DOCUMENT RESUME

ED 347 206 TM 018 737

AUTHOR Shepherd, Jane; Malizio, Andrew G.
TITLE Methodology Report for the 1990 National Postsecondary Student Aid Study. Contractor Report. Technical Report.
INSTITUTION Westat, Inc., Rockville, MD.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
REPORT NO NCES-92-080
PUB DATE May 92
NOTE 286p.
PUB TYPE Statistical Data (110) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC12 Plus Postage.
DESCRIPTORS Cohort Analysis; Enrollment; Higher Education; *Institutional Characteristics; Longitudinal Studies; *National Surveys; Parent Financial Contribution; Paying for College; Postsecondary Education; *Research Methodology; Sampling; *School Surveys; Statistical Data; *Student Characteristics; Student Financial Aid; Tables (Data)
IDENTIFIERS *National Postsecondary Student Aid Study; *Student Surveys

ABSTRACT

The 1990 National Postsecondary Student Aid Study (NPSAS:90) is the second in a series of federally funded surveys of enrolled postsecondary students. Its key purpose is to portray accurately the characteristics of these students, particularly recipients of student aid. Results will answer questions about recipients and sources of federal and private aid and parental support. The NPSAS surveys also provide base information on cohorts that will be the subjects of study over time. Data from the survey can be used for more advanced statistical and economical analyses. The NPSAS:90 used a three-stage sampling technique to select participants (schools and students), and a two-phase data collection method of record abstracting to compile student records from student and parent telephone interviews and data from institutions. In all, 1,130 institutions participated, and 70,200 students were sampled. The executive summary contains 16 tables of student and institutional characteristics, and there are 23 other tables in the body of the report. Twenty-six figures and exhibits illustrate the discussion. Five appendixes present sample material mailed to sampled institutions and students, analysis variables' specifications, procedures used for data imputations, standard errors for estimates in the executive summary, and memos on weighting procedures and a description of the 1987 NPSAS sample. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to: TM

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view. HE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

NATIONAL CENTER FOR EDUCATION STATISTICS

Technical Report

May 1992

Methodology Report for the 1990 National Postsecondary Student Aid Study

Contractor Report

WESTAT, Inc.
Rockville, Maryland
Jane Shepherd, Project Director

Andrew G. Malizio
Project Officer
National Center for Education Statistics

U.S. Department of Education
Office of Educational Research and Improvement

NCES 92-080

U.S. Department of Education
Lamar Alexander
Secretary

Office of Educational Research and Improvement
Diane Ravitch
Assistant Secretary

National Center for Education Statistics
Emerson J. Elliott
Acting Commissioner

National Center for Education Statistics

"The purpose of the Center shall be to collect, and analyze, and disseminate statistics and other data related to education in the United States and in other nations."—Section 406(b) of the General Education Provisions Act, as amended (20 U.S.C. 1221e–1).

May 1992

Contact:
Andrew G. Malizio
(202) 219–1448

ACKNOWLEDGMENTS

The 1990 National Postsecondary Student Aid Study was a large-scale and complex survey of postsecondary students. Tens of thousands of students and parents, thousands of school administrators, hundreds of survey field staff, and dozens of statisticians, analysts, managers, programmers, and other professionals participated in the project. The authors of this report would like to acknowledge gratefully all their efforts.

Researchers at Westat, Inc. of Rockville, MD, conducted this study for the Postsecondary Education Statistics Division's Longitudinal Studies Branch of the National Center for Education Statistics in the U.S. Department of Education. Contributors to this report include: Jane Shepherd, Project Director, J. Michael Brick, Bradford Chaney, Richard Wabnick, Marianne Winglee, Rich Hilpert, Diane Ward, and Kay Sigle. We would like to acknowledge the support and contributions of C. Dennis Carroll, Longitudinal Studies Branch Chief, and Andrew G. Malizio, Project Officer, to this study.

NCES staff who also reviewed this report and provided helpful suggestions include Susan Ahmed, Nabeel A. Alsalam, Roslyn Korb, Sandra Garcia, and Brian Gregory. David Bergeron, Dan Goldenberg and Dan Madzelan from the Department of Education and Laura Greene Knapp from The College Board also provided suggestions for improvement.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

LIST OF TABLES vi
LIST OF FIGURES AND EXHIBITS viii

EXECUTIVE SUMMARY ix

HIGHLIGHTS xix

INTRODUCTION 1

CHAPTER 1. STUDY OBJECTIVES AND DESIGN 3
1.1 Study Objectives 3
1.2 Sample Design 3
1.3 Instrument Design 5
1.4 Data Collection Methods 7

CHAPTER 2. INSTITUTION SAMPLING AND ENLISTMENT 9
2.1 Area Sampling 9
2.2 Institution Sample 10
2.3 Institution Sample Selection 13
2.4 Institution Enlistment 17
2.5 Institution Participation Rates 21

CHAPTER 3. STUDENT SAMPLING 25
3.2 Selection of the Fall and Nonfall Samples of Students 29
3.3 Quality Control and Problem Handling in the Sampling Process 35

CHAPTER 4. INSTITUTIONAL RECORDS DATA COLLECTION 39
4.1 Record Abstracts 39
4.2 Institution Checklist 41
4.3 Scheduling Institution Visits 42
4.4 Field Data Collection 49
4.5 Supervisor Recruitment and Training 53
4.6 Data Collector Recruitment and Training 55
4.7 Field Visits 62
4.8 Receipt and Processing 64
4.9 Nonfall Records Data Collection 65

CHAPTER 5. INSTITUTIONAL RECORDS UPDATE 67
5.1 Records Update Form 67
5.2 The Control List 67
5.3 The Other Specify Awards Form 67
5.4 Data collection 71
5.5 Receipt Control 73

CHAPTER 6: STUDENT SURVEY **75**
6.1 Objectives 75
6.2 Instrument Design 75
6.3 Staff Organization 77
6.4 Interviewer Training 78
6.5 Data Collection 78
6.6 Locating Students 84
6.7 Language Problems 84
6.8 Toll-Free 800 Number 86
6.9 Quality Control 86
6.10 Reports 86
6.11 Results of Data Collection 86
6.12 Data Preparation 90

CHAPTER 7: PARENT SURVEY **93**
7.1 Objectives 93
7.2 Instrument Design 93
7.3 Staff Organization 94
7.4 Interviewer Training 94
7.5 Data Collection 95
7.6 Locating Parents/Guardians 99
7.7 Language Problems 99
7.8 Toll-Free 800 Number 100
7.9 Quality Control 100
7.10 Reports 100
7.11 Results of Data Collection 101
7.12 Data Preparation 101

CHAPTER 8: FILE CREATION AND DATA ANALYSIS **107**
8.1 Editing 107
8.2 Coding 110
8.3 The NPSAS Analysis File 113

CHAPTER 9. WEIGHTS AND VARIANCE ESTIMATION **123**
9.1 Weighting 123
9.2 Variance Estimation 132
9.3 Coverage Issues 136

CHAPTER 10. SUMMARY OF THE 1990 NPSAS FIELD TEST **147**

APPENDICES
Appendix A. Sample Materials Mailed to Sampled Institutions and Students A-1
Appendix B. Analysis Variables Specifications B-1
Appendix C. Procedures Used for Data Imputations C-1
Appendix D. Standard Errors for Estimates in the Executive Summary D-1
Appendix E. Memos on Weighting Procedures and Description of 1987 NPSAS sample . . E-1

LIST OF TABLES

Tables in Executive Summary

A Institutional response rates: Number and percentage of institutions participating in NPSAS:90 xi

B Expected student sample for NPSAS:90: Sample size for the full-year by institution sector and student educational level xii

C Student Sample for NPSAS:90: Number of Eligible Students Sampled in Fall and Nonfall and by Institution Level, Student Education Level xiv

1 Number of Students by Type and Control and Academic Level: 1989-90 xxiii

2 Number of Students, by Family Income and Academic Level: 1989-90 xxiii

3 Number of Students, by Type and Control and Dependency Status: 1989-90 xxiv

4 Number of Students, by Family Income and Dependency Status: 1989-90 xxiv

5 Percentage of Students Receiving Title IV Aid and Any Aid, by Type and Control and Academic Level: 1989-90 xxv

6 Percentage of Students Receiving Title IV Aid and Any Aid, by Family Income and Academic Level: 1989-90 xxv

7 Percentage of Students Receiving Title IV Aid and Any Aid, by Type and Control and Dependency Status: 1989-90 xxvi

8 Percentage of Students Receiving Title IV Aid and Any Aid, by Family Income and Dependency Status: 1989-90 xxvi

9 Average Title IV and Total Aid of Students, by Type and Control and Academic Level: 1989-90 xxvii

10 Average Title IV and Total Aid of Students, by Family Income and Academic Level: 1989-90 xxvii

11 Average Title IV and Total Aid of Students, by Type and Control and Dependency Status: 1989-90 xxviii

12 Average Title IV and Total Aid of Students, by Family Income and Dependency Status: 1989-90 xxviii

LIST OF TABLES (continued)

Tables in Report

2.2.1 Frame Development: Number of institutions in the 1987-88 IPEDS-IC File and in the NPSAS:90 institutional frame 12
2.3.1 Institution sample: Number of institutions in the frame and in the sample 15
2.3.2 Institution sample: Number of self-representing and non-self-representing institutions from the entire frame, self-representing NPSAS PSUs and non-self-representing NPSAS PSUs 16
2.3.3 Institution sample: Enrollment cutoff for self-representing institutions from entire frame and sampling intervals for institutions in NPSAS PSUs 17
2.4.1 Source of Institution Enlistment Status for Participating Institutions 22
2.5.1 Institutional response rates: Number and percentage of institutions participating in NPSAS:90 23

3.1.1 Expected student sample for NPSAS:90: Sample size for the full-year by institutional sector and student educational level 28
3.2.1 Student sample for NPSAS:90: Sample size for the fall sample by institutional sector and student educational level 32
3.2.2 Student sample for NPSAS:90: Estimated population size for fall enrollment by institutional sector and student educational level 33
3.2.3 Student sample for NPSAS:90: Expected number of students for the nonfall sample by institutional sector and student educational level 34
3.2.4 Student sample for NPSAS:90: Sampling time points for groups of institutions 35

8.1 1990 NPSAS data files summary table 108
8.2 NPSAS Analysis Variables 115
8.3 Summary statistics on imputations for the NPSAS:90 derived variable file 122

9.1.1 Ratio adjustment for abstract estimates based on Pell counts, by level and control of institution and Pell grant award amount 130
9.2.1 Ratio adjustment for student analysis file based on Pell grant counts, by level and control of institution and Pell grant award amount 130
9.3.1 Estimated percentage of students enrolled in more than one postsecondary institution during 1989-90 131
9.4.1 Estimates and standard errors for categorical data in NPSAS 135
9.5.1 Design effects for categorical data in NPSAS 136
9.6.1 Ratios of the estimated weighted number of undergraduates in NPSAS eligible institutions to the number of undergraduates as reported in IPEDS 140
9.7.1 Ratio of NPSAS estimates to Pell grant program totals, by institution level and control 142
9.8.1 Ratios of the Pell grant estimates from NPSAS participating institutions to the Pell grant universe 143

10.1 Schools in Pretest 150

LIST OF FIGURES AND EXHIBITS

Figure		
2.2.1	Institutions Eligible for NPSAS:90	11
3.1.1	Students eligible for NPSAS:90	27
4.1	Data Collection Scheduling Information	44
4.2	Sample Scheduling Calendar	46
4.3	Field Group Assignment Summary Sheet	48
4.4	Data Collector's Task	50
4.5	Sample Disclosure Notice	52
4.6	Field Data Collector Training Manual Table of Contents	58
5.1	Records Update Form	68
6.1	Student Population	76
6.2	Number of Students Sampled, Number Eligible, and Number Included the Analysis File for NPSAS:90	80
6.3	Number of eligible students included in NPSAS:9O analysis file	81
6.4	Distribution of completed cases by day and time.	82
6.5	Cumulative number of cases traced	85
6.6	Student CATI response rates, by student characteristics	87
6.7	Percentage of completed cases by data collection month	89
6.8	CATI administration time, by section	89
6.9	Student survey file structure	91
7.1	Parent Interview, distribution of completed cases by time of day	102
7.2	Parent Interview, cumulative number of cases traced	103
7.3	Parent Interview, average completion times by section	104
7.4	Number of Parents Selected for NPSAS:9O Parent Survey, Number Eligible, and Number Participated	105
7.5	Parent Interview file structure	106
8.1	Derivation of Student Costs	118
10.1	NPSAS:90 Field Test Results	148
10.2	NPSAS:90 Field Test Design	149

Executive Summary
of the
1990 National Postsecondary Student Aid Study
Methodology Report

Study Purposes

The 1990 National Postsecondary Student Aid Study (NPSAS:90) is the second in a series of federally-funded surveys of enrolled postsecondary students. Its key purpose is to portray accurately the characteristics of those students, particularly student aid recipients. Results from NPSAS:90 will answer such basic questions as: What percentage of students receive federal student aid and in what amounts? What percentage receive state aid and in what amounts? What percentage receive institutional aid and in what amounts? What percentage receive other private assistance? What forms of aid do students receive--grants, loans, or work-study? How much do parents contribute to their children's education? This series of studies is the only national-level source for answers to these and other questions about postsecondary students, their schools, their programs of study, and their financial aid.

Beyond this descriptive purpose, the NPSAS surveys serve two other important purposes. First, they provide the base information for special groups (or cohorts) of students who will be the subject of study over time. NPSAS:90 included a longitudinal cohort of first-time students. NPSAS:93 will include a longitudinal cohort for graduating seniors. Second, researchers and analysts can use the data collected through these surveys for more advanced statistical and economic analyses. For example, policy analysts may use the data to measure the effects of aid changes on enrolled students.

In addition, policymakers and their advisors will no doubt apply the NPSAS data to many fundamental questions about the federal role in financing postsecondary education. Once enrolled, are students prevented from completing their course of study by financial or other barriers? NPSAS data, though limited to the enrolled population, can be used to address these questions.

Study Methods

NPSAS:90 was a complex sample survey. It used a three-stage sampling technique to select participants (schools and students). It also used a two-phase data collection method of record abstracting to compile student and parent data from institutions, and student and parent telephone interviews.

Sample Design

The NPSAS:90 sampling design was a multi-stage probability sample students enrolled in postsecondary institutions at any time between July 1, 1989--June 30, 1990. The first stage of sampling was the selection of geographic areas of the country, called primary sampling units (PSUs). For NPSAS:90, 121 PSUs were selected (For detail on PSU selection, see Section 2.1). The second stage of the sample design was the sampling of institutions in the NPSAS PSUs. Institutions were stratified by control, type and enrollment--i.e., whether the institution enrolled at least 75 students in first-professional, doctoral, master's, or baccalaureate degree programs (see Table 2.3.1 for original sampling stratum). Then, institutions were sampled with probabilities proportionate to size in each stratum. In all, 1,533 institutions were sampled. About 80 percent of the sampled schools were eligible and, of these, 1,130 participated. (Table A displays institutional participation rates by type and control.) The third stage of sampling was the selection of students in sampled institutions. Because one of the goals of NPSAS:90 was to estimate full-year award information, four student samples were drawn across the award year--in August, October, February, June. More than 70,000 students were sampled. (Table B displays the student sample in each sample period by institution type and student level.)

Although four samples were drawn during the year, Westat used a cross-checking method to assure that each student had only one chance of selection from each institution. First, students sampled other than in October (nonfall) were checked against the October sampling list. If a sampled non-fall student was on the October sampling list that student was eliminated from the sample. For example, students sampled from the August list were checked against the entire sampling list for October. Those found on the October list were deleted from the August sample. Likewise, students sampled in February were checked against the October and the August lists. Those sampled in June were checked against the August, October, and February lists. Students found in any prior sampling lists were deleted. This checking was also done across educational levels to assure that any student who changed education level during the year would still have only one chance of selection.

Table A.--Institutional response rates: Number and percentage of institutions participating in NPSAS:90

Institutional Sector	Number of:		Response Rates	
	Respondents	Non-Respondents	Unweighted	Weighted
Public, less-than-2-year	62	3	95%	89%
Public, 2-year	189	16	92%	96%
Public, other 4-year	113	7	94%	92%
Public, doctoral	115	9	93%	93%
Private, not-for-profit, less-than-2-year	34	3	92%	93%
Private, not-for-profit, 2-year	59	10	86%	86%
Private, not-for-profit, other 4-year	146	13	92%	88%
Private, not-for-profit, doctoral	128	12	91%	91%
Private, for-profit, less-than-2-year	206	37	85%	80%
Private, for-profit, 2-year or more	78	8	91%	87%
Certainty institutions	259	26	91%	91%
Non-certainty institutions	871	92	90%	86%
Total	1,130	118	91%	86%

Note: Chapter 9 of the Methodology Report describes weighting and variance estimation.

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

BEST COPY AVAILABLE

Table B. Expected student sample for NPSAS:90: Sample size for the full-year by institution sector and student educational level

Institutional Sector	Number of Students			
	Total	Undergraduate	Graduate	First-Professional
Public, doctoral	12,886	7,036	3,650	2,200
Private, doctoral	15,933	9,543	2,590	3,800
Public, other 4-year	8,791	6,911	1,880	
Private, other 4-year	9,590	7,710	1,880	
Public, 2-year	7,530	7,530		
Private, not-for-profit, 2-year	2,259	2,259		
Private, for-profit, 2-year	2,710	2,710		
Public, less than 2-year	1,800	1,800		
Private, not-for-profit, less-than-2-year	1,500	1,500		
Private, for-profit, less-than-2-year	7,201	7,201		
Total Sampled	70,200	54,200	10,000	6,000

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

Interviewers contacted students in the NPSAS:90 sample by telephone and those who satisfied the eligibility criteria were asked to participate. The number of eligible students from the fall and the non-fall samples are shown in table C.

Data Collection

NPSAS:90 included three sources of data: institutional records, student interviews, and parent interviews. The basic statistics about these sources are summarized below:

Institutional Records Data

- on-campus data abstraction from about 70,000 institutional records at 1,130 schools
- end-of-year updates of institutional records at 488 schools
- 158 record abstractors, average time to complete each school's data abstracts was 2.5 days
- 46 update abstractors, average time to complete each school's updates was 2 days
- 2 schools self-reported on the initial data collection

Student Survey

- 479 interviewers completed over 51,000 computer-assisted telephone student surveys
- approximately 540,000 calls were required
- average time to complete a student interview was 45 minutes
- interviewers achieved a weighted response rate of 76 percent

Table C. Student Sample for NPSAS:90: Number of Eligible Students Sampled in Fall and Nonfall and by Institution Level, Student Education Level

Institutional Level and Student Level	Students Sampled		
	Total	Fall Samples	Nonfall Samples
Total	70,000	57,000	13,000
Less-than-2-year institutions	10,500	5,500	5,000
2-year institutions	12,500	8,500	4,000
4-year and above institutions	47,000	43,000	4,000
Undergraduate students	31,000	29,000	2,000
First-professional students	6,000	5,500	500
Graduate students	10,000	8,500	1,500

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

Parent Survey

- about 150 interviewers completed more than 16,000 computer-assisted telephone parent surveys
- more than 100,000 calls were required
- average time to complete a parent interview was 20 minutes

Trained data collectors visited institutions and used a special **record abstract** to collect the data. Data collection was conducted during two periods: December 1989 through February 1990 and August 1990 through November 1990.

A designated coordinator at each institution provided information on the best sources and locations for the requested data. The record abstract was developed to standardize the collection of those data. It contained sections on:

- student enrollment status and locating information
- financial aid award information, by source (federal, state, institution, private)
- need analysis and budget information, including the Student Aid Index and Expected Family Contributions, and allowable costs
- financial aid application data from, for example, the federal Student Aid Report (SAR) or the CSS Financial Aid Form Need Analysis Report (FAFNAR)

The initial records data collection was followed up by a **records update** later in the school year. This updating identified any changes in financial aid award amounts or in any variables which could affect award amounts, such as dependency or enrollment status.

The **student survey** was conducted by telephone. Westat's computer-assisted telephone interviewing system was used because it allowed on-line editing of data during questionnaire administration. It had the added advantage of automatically adjusting for complex skip patterns in the questionnaire.

The student survey instrument contained the following eleven sections:

- introduction
- school enrollment
- enrollment and costs
- financial aid
- additional sources of support
- employment
- educational expectations/student characteristics
- parental characteristics
- financial status
- demographics
- locating information

Like the student survey, the **parent survey** was also conducted using a computer-assisted telephone interview system. The parent survey, however, was a purposive subsample drawn from the student survey. It was directed primarily at parents of undergraduate, dependent, unaided students. Parents were asked to describe what financial support they had provided to the sampled student and about other dependents whom they supported. Further, parents were asked to describe their own financial condition, including their income, type of job, monthly expenses, assets, and the funds borrowed for education purposes.

Coverage Issues

The extent to which the population of inference is represented or covered in the sampling frame is an important concern. In surveys of postsecondary institutions and students, undercoverage can occur for a variety of reasons:

- the frame from which the sample of institutions was selected may exclude some types of institutions,
- the institutions which have recently been established may be excluded,
- the list of students provided by the institutions may not include off-campus students, and
- the list of nonfall students may be matched incorrectly against the list of previously enrolled students.

For NPSAS:90, we concluded there is no evidence of a large bias associated with undercoverage of institutions. The potential for undercoverage in estimates due to the enrollment lists not including all students needs further investigation. Some unduplicated headcounts of students for many postsecondary institutions are available using IPEDS and other Office of Postsecondary Education data files. However, sufficient quantitative evidence to examine fully the student coverage within all institutions does not exist. To evaluate the student coverage within institutions, an accurate unduplicated headcount of enrolled students for each postsecondary institution is needed.

File Preparation

On completion of data collection, the preparation of four basic data files began. Three raw data files were created which contain the results from the institution records', students', and parents' surveys. These were then edited and coded for consistency and accuracy. Award data on these files are for each award period and for each school in which a student is enrolled. The fourth file is an **analysis file** which contains a selected set of variables derived from the raw data files. Awards on this analysis file are all full-year awards. This analysis file, which contains 300 variables and 300 matching source flags, is intended as the primary research file for the NPSAS:90. Following is a brief description of each file and a chart summarizing some pertinent file characteristics.

Summary Table
1990 NPSAS Data Files

File	Respondents	Maximum Number of Records	Key Elements	Number of Variables
1. Institutional Records Data	Institutions	68,599**	Demographic, Budget and Application Data	431
2. Institutional Awards Data	Institutions	403,477**	Dates, Types and amounts of awards	32
3. Student Survey Data	Postsecondary students	69,613**	Demographic, employment and financial aid data	498
4. Parent Survey Data	Parents of postsecondary students	16,106	Parent demographics, sources of support, and attitudes	241
5. Analysis File	Integrated Institution, Student and Parent	61,120	Full-year award amounts, demographics, budget and application data	665

Note: **Institutional records data, in addition to locating information were collected for 68,599 students. Telephone interview data were collected for 51,430 of the 66,718 (NPSAS CATI-eligible) students. However, the institutional awards data file and some of the student survey data files include multiple records per student. The student survey data jobs file contains one record for each job--a total of 69,613 records. The institutional awards data file contains one record for each award received or institutional budget data element. Some students may have up to 32 records in this file, which contains nearly 403,500 records.

Data Access

The data collected in NPSAS:90 is stored on a CD-ROM disk. One disk, which is similar in size to an audio CD, can store large quantities of data (over 600 Megabytes) in an extremely compact space. The NPSAS:90 CD-ROM contains all data from the raw and analysis files and an electronic codebook. The codebook allows the data user to view frequency distributions of the responses to each survey item as well as to create extract files with which to perform other analyses. Those seeking access to the NPSAS:90 CD-ROM data must obtain a license agreement from the National Center for Education Statistics, statistical Standards and Methodology Division.

NCES has developed a table generation system which permits users to produce tables the using variables from the NPSAS:90 analysis file (for example, a table showing the average amount of aid received by type of institution attended). Information on the NPSAS:90 table generation system is available from the Longitudinal Studies Branch.

Major Findings

The major descriptive findings of NPSAS:90 are summarized in the charts presented at the end of this executive summary (for standard errors, see Appendix D). These charts present the data across five key variables: aid status (aid or no aid), institutional type and control (public 4-year, public 2-year, private 4-year, private 2-year, and proprietary), income (in categories up to $100,000), academic level (undergraduate/graduate/first-professional), and dependency status (dependent or independent, as defined for aid purposes).

Because these statistics come from a sample survey, they may differ from figures from a complete census of the same population. A sample survey may contain two types of error: sampling and nonsampling. The accuracy of a survey statistic depends on both types of error, but the full extent of nonsampling error is generally unknown. Sampling variability is variation that occurred by chance because a sample was surveyed rather than the population. This variation is commonly represented by the standard error of the estimate. Sources of nonsampling errors include nonresponse, misinterpretation, coding errors, processing errors, and coverage errors.

Highlights from NPSAS:90 (from a recent NCES Tabulation #92-003)

- About 18.6 million students were enrolled in 1989-90--including 16.3 million undergraduates and 2.3 million graduate and first-professional students.

Among the 16.3 million **undergraduates** (including full-time and part-time students) enrolled during 1989-90:

- About 43 percent (nearly 7 million) received financial aid from any source, including federal or state governments, institutions, or other private organizations, or combinations of these sources (excluding aid from relatives); averaging about $3,600.

- About 3 of every 10 received some type of federal aid; about 2 of every 10 received federal grants.

- Percentages of students receiving financial aid varied considerably, depending on the type of institution. Percentages ranged from about 28 percent of the 6.8 million undergraduates at public 2-year institutions to 82 percent of the 900,000 enrolled at proprietary, less-than-2-year schools.

- Average amounts of aid also varied considerably, depending on the type of institution.

  Among the 1.9 million aided students enrolled at public 2-year schools, the average amount of aid received was about $2,000.

  Among the 500,000 aided students enrolled in private, not-for-profit, doctoral-granting institutions, the average amount received was about $7,100.

  For the 750,000 aided students enrolled in private, for-profit, less-than-2-year institutions, the average amount was $4,100.

  Among the 1.3 million aided students enrolled in public, doctoral institutions, the average amount was about $3,600.

  Among the 985,000 aided students enrolled in other public 4-year institutions, the average amount was about $3,000.

- Overall, about 36 percent received some grant aid (including grants from federal and state governments, institutions, and/or employers). More than 70 percent of dependent undergraduates from families with incomes less than $10,000 received some grant aid, averaging about $2,900.

- Overall, about 10 percent of undergraduates in private, 4-year, not-for-profit institutions and 4 percent in public, 4-year institutions received federal college work-study assistance, averaging about $1,000.

- Among undergraduates, 29 percent received Title IV aid, whether dependent or independent. The average amount of Title IV aid to dependent students was $2,924 while to independent students it was $3,107.

Among the 2.3 million **graduate and first-professional** students (including full-time and part-time students) enrolled during 1989-90:

- About 45 percent (1 million) received some financial aid from any source, including federal or state governments, institutions, or employers; averaging nearly $8,000.

- About 18 percent received some type of federal aid, averaging $8,100; about 1 of every 4 received some institutional aid, averaging about $6,700; 1 of every 10 received some

employer assistance, averaging about $1,900.

- Percentages of students receiving financial aid varied considerably, depending on the type of degree program. Percentages ranged from about 40 percent of the 1.3 million students enrolled in master's programs to about 70 percent of the 300,000 students enrolled in first-professional programs (e.g., law school, medical school, dentistry).

- Average amounts varied considerably, depending on the type of program. Among the 535,000 aided students in master's programs, the average amount of aid received was $5,900. For the 150,000 aided doctoral students and the 215,000 aided first-professional students, the average amount was more than $12,000.

- Overall, about 30 percent received some grant aid (including grants from federal and state governments, institutions, and/or employers), averaging about $3,200.

- Overall, nearly 3 of every 10 doctoral students received some type of assistantship, averaging about $9,000.

- Among graduate students, 17 percent received Title IV aid and the average amount of Title IV aid was $7,275.

Operational Recommendations

In this section, we outline six operational concerns we believe would increase the efficiency of future NPSAS surveys. We have based these concerns on our knowledge of recent technological advances and our experience with NPSAS survey methods.

Technological innovations. Several technological innovations entered the survey world during the 1990 NPSAS. The most important was the development of lightweight but powerful laptop computers. These smaller machines spurred the development of computer-assisted data entry since in-person interviewers or data abstractors could take preprogrammed survey instruments into the field. The survey software built into these machines can contain full editing and coding specifications so the resulting data are clean and ready for final file preparation. This preloaded survey capability has several advantages, including, reducing manual editing after data collection and accelerating the file preparation process. Since these new laptops also contain modems, data can be transmitted electronically between field stations and a central location. As a result, file creation becomes a continuous rather than a sequential process.

Another recent technological improvement was the enhanced networking ability of personal computers. Besides passing data files from one machine to another, communications packages are now available which permit multiple users to access text documents, to make revisions to those documents and to transmit those revisions immediately to each user.

Another advance was the development of a more economical means for storing data on compact discs and replicating those discs. This new storage medium, called a CD-ROM, will be used to disseminate 1990 NPSAS findings. All of the NPSAS:90 files, including the electronic codebook are stored on one compact disk (CD-ROM). By contrast, more than 300 high-density floppy disks would be required to store the NPSAS:90 data.

Suggestions for improving overall operations. Our suggestions are as follows:

- Thoroughly **evaluate and pretest** sampling methods and survey instruments to simplify reporting requirements and reduce institutional, student, and parent response burden.

- **Enhance training** for field data collectors using a fully scripted approach and role-playing.

- Install a **computer-assisted data entry system** for use by field staff which will permit:
  - -- **preloading** of data to avoid duplicate requests
  - -- upfront editing checks to **reduce post data collection editing**
  - -- **flexible interviewing** of first respondent contacted.

- **Improve communications links** among project staff, field staff, and NCES.

- Apply **latest technology for data storage, retrieval, and tabulation**, including data libraries on CD ROMs, encrypted data tabulation systems, electronic codebooks.

- **Integrate survey operations** (communications, forms clearance, data entry, editing, and file preparation) to create a more efficient processing and file generation system.

Summary Table 1

Number of Students by Type and Control and Academic Level, in Thousands: 1989-90

Academic Level	All Institutions	Institutional Type and Control						
		Undergraduate					Graduate	
		Public Four-year	Public Two-year	Private Four-year	Private Two-year	Proprietary	Public	Private
Undergraduate	16,271	5,260	7,052	2,298	269	1,391	--	--
Graduate	2,318	--	--	--	--	--	1,440	879
All	18,590	5,260	7,052	2,298	269	1,391	1,440	879

Summary Table 2

Number of Students, by Family Income and Academic Level, in Thousands: 1989-90

Academic Level	All Incomes	Family Income (adjusted gross income)						
		Less than $10,000	$10,000-$20,000	$20,000-$30,000	$30,000-$40,000	$40,000-$50,000	$50,000-$100,000	$100,000 and over
Undergraduate	16,271	4,023	3,072	2,569	2,106	1,599	2,514	389
Graduate	2,318	596	443	417	307	220	318	16
All	18,590	4,619	3,516	2,986	2,413	1,819	2,832	405

Summary Table 3

Number of Students, by Type and Control and Dependency Status, in Thousands: 1989-90

Students	All Institutions*	Institutional Type and Control						
		Undergraduate					Graduate	
		Public Four-year	Public Two-year	Private Four-year	Private Two-year	Proprietary	Public	Private
Dependent	7,846	3,345	2,402	1,503	117	391	52	35
Independent	10,679	1,900	4,633	783	148	983	1,388	843
All*	18,590	5,245	7,052	2,286	269	1,391	1,400	878

Summary Table 4

Number of Students, by Family Income and Dependency Status, in Thousands: 1989-90

Students	All Incomes*	Family Income (adjusted gross income)						
		Less than $10,000	$10,000-$20,000	$20,000-$30,000	$30,000-$40,000	$40,000-$50,000	$50,000-$100,000	$100,000 and over
Dependent	7,846	1,006	935	1,148	1,246	1,062	2,064	385
Independent	10,679	3,548	2,581	1,837	1,167	758	768	20
All*	18,590	4,619	3,516	2,986	2,413	1,819	2,832	405

* Since 65,500 weighted cases were unclassified, numbers do not add to totals.

Summary Table 5

Percentage of Students Receiving Title IV Aid and Any Aid, by Type and Control and Academic Level: 1989-90

Academic Level		All Institutions	Institutional Type and Control						
			Undergraduate					Graduate	
			Public Four-year	Public Two-year	Private Four-year	Private Two-year	Proprietary	Public	Private
Undergraduate	percent Title IV aid	28.8	33.2	22.4	20.2	2.2	22.1	--	--
	percent any aid	42.8	32.5	28.3	20.9	2.3	16.1	--	--
Graduate	percent Title IV aid	17.1	--	--	--	--	--	46.9	53.1
	percent any aid	45.1	--	--	--	--	--	54.5	45.5
All	percent Title IV aid	27.3	30.6	20.6	18.6	2.0	20.4	3.7	4.1
	percent any aid	43.1	28.2	24.6	18.2	2.0	14.0	7.1	5.9

Summary Table 6

Percentage of Students Receiving Title IV Aid and Any Aid, by Family Income and Academic Level: 1989-90

Academic Level		All Incomes	Family Income						
			Less than $10,000	$10,000-$20,000	$20,000-$30,000	$30,000-$40,000	$40,000-$50,000	$50,000-$100,000	$100,000 and over
Undergraduate	percent Title IV aid	28.8	45.0	21.0	14.1	8.8	5.3	5.5	0.3
	percent any aid	42.8	36.0	19.5	15.1	11.0	7.7	10.0	0.9
Graduate	percent Title IV aid	17.1	48.5	21.5	13.7	7.2	3.9	5.0	0.2
	percent any aid	45.1	35.1	20.6	16.3	10.4	7.1	10.0	0.5
All	percent Title IV aid	27.3	45.3	21.1	14.1	8.7	5.2	5.5	0.3
	percent any aid	43.1	35.9	19.6	15.2	10.9	7.6	10.0	0.8

Summary Table 7

Percentage of Students Receiving Title IV Aid and Any Aid, by Type and Control and Dependency Status: 1989-90

Students		All Institutions	Institutional Type and Control						
			Undergraduate					Graduate	
			Public Four-year	Public Two-year	Private Four-year	Private Two-year	Proprietary	Public	Private
Dependent	percent Title IV aid	28.7	41.4	14.7	29.5	2.0	11.4	0.4	0.6
	percent any aid	42.9	41.7	17.3	29.3	1.9	8.3	0.8	0.6
Independent	percent Title IV aid	26.4	22.0	25.4	9.9	2.1	27.4	6.2	7.0
	percent any aid	43.2	18.5	29.9	10.0	2.0	18.0	11.8	9.8
All	percent Title IV aid	27.3	30.6	20.6	18.6	2.0	20.4	3.7	4.1
	percent any aid	43.1	28.2	24.6	18.2	2.0	14.0	7.1	5.9

Summary Table 8

Percentage of Students Receiving Title IV Aid and Any Aid, by Family Income and Dependency Status: 1989-90

Students		All Incomes	Family Income (adjusted gross income)						
			Less than $10,000	$10,000-$20,000	$20,000-$30,000	$30,000-$40,000	$40,000-$50,000	$50,000-$100,000	$100,000 and over
Dependent	percent Title IV aid	28.7	28.5	18.2	18.0	14.5	9.5	10.8	0.5
	percent any aid	42.9	21.9	15.0	16.5	15.5	12.1	17.2	1.8
Independent	percent Title IV aid	26.4	58.5	23.4	11.1	4.1	1.7	1.2	0.1
	percent any aid	43.2	45.5	23.1	14.5	7.6	4.3	4.8	0.1
All	percent Title IV aid	27.3	45.3	21.1	14.1	8.7	5.2	5.5	0.3
	percent any aid	43.1	35.9	19.6	15.2	10.9	7.6	10.0	0.8

Summary Table 9

Average Title IV and Total Aid of Students, by Type and Control and Academic Level: 1989-90

Academic Level		All Institutions	Institutional Type and Control						
			Undergraduate					Graduate	
			Public Four-year	Public Two-year	Private Four-year	Private Two-year	Proprietary	Public	Private
Undergraduate	average Title IV aid	$3,020	$2,978	$1,847	$3,773	$2,876	$3,593	--	--
	average total aid	3,606	3,351	1,991	5,845	3,471	4,066	--	--
Graduate	average Title IV aid	7,275	--	--	--	--	--	$6,158	$8,260
	average total aid	7,987	--	--	--	--	--	6,755	9,465
All	average Title IV aid	3,351	2,978	1,847	3,773	2,876	3,593	6,158	8,260
	average total aid	4,177	3,351	1,991	5,845	3,471	4,066	6,755	9,465

Summary Table 10

Average Title IV and Total Aid of Students, by Family Income and Academic Level: 1989-90

Students		All Incomes	Family Income (adjusted gross income)						
			Less than $10,000	$10,000-$20,000	$20,000-$30,000	$30,000-$40,000	$40,000-$50,000	$50,000-$100,000	$100,000 and over
Undergraduate	average Title IV aid	$3,020	$3,122	$3,023	$2,855	$2,807	$2,895	$3,034	$3,329
	average total aid	3,606	3,923	3,482	3,306	3,437	3,456	3,434	3,881
Graduate	average Title IV aid	7,275	7,554	7,115	6,877	7,019	6,750	7,127	7,215
	average total aid	7,987	10,558	9,001	5,965	5,799	4,618	4,777	9,548
All	average Title IV aid	3,351	3,492	3,349	3,158	3,081	3,124	3,323	3,591
	average total aid	4,177	4,769	4,237	3,678	3,730	3,597	3,609	4,344

Note: For tables 9-12, average Title IV aid based on Title IV aid recipients only. Average total aid based on aided students only.

Summary Table 11

Average Title IV and Total Aid of Students, by Type and Control and Dependency Status: 1989-90

Students		All Institutions	Institutional Type and Control						
			Undergraduate					Graduate	
			Public Four-year	Public Two-year	Private Four-year	Private Two-year	Proprietary	Public	Private
Dependent	average Title IV aid	$2,971	$2,709	$1,500	$3,738	$2,731	$3,472	$5,977	$ 8,743
	average total aid	4,166	3,257	2,159	6,498	3,527	3,943	6,596	12,756
Independent	average Title IV aid	3,656	3,383	2,008	3,859	2,987	3,639	6,168	8,228
	average total aid	4,205	3,524	1,931	4,524	3,462	4,123	6,763	9,310
All	average Title IV aid	3,351	2,978	1,847	3,773	2,876	3,593	6,158	8,260
	average total aid	4,177	3,351	1,991	5,845	3,471	4,060	6,755	9,465

Summary Table 12

Average Title IV and Total Aid of Students, by Family Income and Dependency Status: 1989-90

Students		All Incomes	Family Income (adjusted gross income)						
			Less than $10,000	$10,000-$20,000	$20,000-$30,000	$30,000-$40,000	$40,000-$50,000	$50,000-$100,000	$100,000 and over
Dependent	average Title IV aid	$2,971	$2,880	$3,120	$2,920	$2,850	$2,972	$3,178	$3,550
	average total aid	4,166	4,311	4,417	4,096	4,158	3,914	3,994	4,392
Independent	average Title IV aid	3,656	3,732	3,492	3,467	3,736	3,809	4,350	low N
	average total aid	4,205	4,984	4,152	3,331	3,097	2,945	2,607	3,888
All	average Title IV aid	3,351	3,492	3,349	3,158	3,081	3,124	3,323	3,591
	average total aid	4,177	4,769	4,237	3,678	3,730	3,597	3,609	4,344

Note: For tables 9-12, average Title IV aid based on Title IV aid recipients only. Average total aid based on aided students only.

INTRODUCTION

This report describes the operations undertaken to complete the 1990 National Postsecondary Student Aid Study (NPSAS:90). It is intended for those who wish to understand the way in which these data were collected, compiled, and analyzed.

While it is detailed, this report does not cover every aspect of the survey at the same level of detail. Where useful we relied on charts and tables for summarizing activities and findings. Overall, we divided the survey operations into chapters which roughly follow the time sequence in which these activities occurred, from study objectives and design to file creation, analysis, and recommendations.

NPSAS:90 contains important changes from the previous NPSAS survey (conducted in 1987) in its sampling design and its collection of data by term. Still, maintaining comparability between the two surveys has remained an important goal, and has guided both questionnaire design and sampling.

Like NPSAS:87, NPSAS:90 was comprehensive in its scope. It included public and private institutions, not-for-profit and for-profit schools. The programs surveyed ranged from short-term occupational to baccalaureate and doctoral programs. The students selected for the sample included undergraduates, graduates, and first-professional students.

The survey consisted of several linked components. First, data were collected from institutional records on major field of study, attendance status, demographic characteristics, financial aid awards, and budget and application data used to determine financial aid amounts. Next, interviewers telephoned sampled students to update and confirm some of the demographic and financial aid data collected from institutions, to collect information on education-related and other costs, and methods of financing postsecondary education costs. Interviewers asked a subsample of about 16,100 students' parents (primarily those of unaided, dependent students) about family finances and other family characteristics. Institutional record data were updated in spring of 1990 to provide a more complete picture of the entire academic year.

Chapter 1 of this report covers the study's central objectives and basic design features. Chapter 2 describes the institutional sampling and enlistment process. Chapter 3 reviews the student sampling procedures. Chapter 4 describes the institutional records data collection and

Chapter 5 reviews the updating of those records. Chapters 6 and 7, respectively, describe the methods used in the student and parent telephone interviews. Chapter 8 reviews the process of creating the raw data files and the analysis file. Chapter 9 describes the computation of sampling weights and standard errors. Chapter 10 summarizes the results of the 1990 Field Test.

CHAPTER 1. STUDY OBJECTIVES AND DESIGN

1.1 Study Objectives

The 1990 NPSAS had three fundamental objectives:

- To collect data to determine how students and their families pay for postsecondary education.
- To estimate percentages of students who received federal student aid, particularly Title IV assistance by type and control of the institution (e.g., public, 4-year; private, 4-year, etc.).
- To create a representative cohort of first-year students who could subsequently be tracked through postsecondary education and beyond.

1.2 Sample Design

Target Population

The target population of NPSAS:90 was students enrolled in postsecondary institutions in the United States and Puerto Rico during the academic year 1989-90. A probability sample of students was selected to represent this population and data were collected from the students, the postsecondary institutions they attended, and a sample of the parents of the participating students.

Survey Frame

Prior to sample selection, a survey frame for NPSAS:90 was developed. The primary data file used to build the frame was the 1987-88 Institution Characteristics (IC) file from the Integrated Postsecondary Education Date System (IPEDS). This file was augmented with institutions from the 1987 IPEDS Fall Enrollment (EF) file, and the 1987-88 Pell Grant institution file. An unduplicated list of institutions from these sources formed the universe for selecting the sample of postsecondary institutions.

Because of the concern that an institutional frame based primarily on IPEDS data may underrepresent small institutions, a list of institutions participating in Stafford Loan programs was compiled. This was used to draw a supplemental sample primarily of small institutions. More details about frame development are discussed in section 2.2.

Sampling Units and Selection

The NPSAS:90 sample was a stratified multi-stage probability sample of students enrolled in postsecondary institutions. The three stages of the sample design were the sampling of areas, called the primary sampling units (PSUs); the sampling of institutions within the sampled PSU's; and the sampling of students within the sampled institutions. Sampling strata were formed by classifying institutions by type, control, enrollment in various degree programs--e.g., first-professional, doctoral, master's, and baccalaureate (see Table 2.3.1); then by classifying students by their educational level (undergraduates, graduates, and first-professionals). NPSAS:90 was designed so that reliable national estimates can be reported for students in each of these domains.

The PSU sample selected for NPSAS:87 was used again in NPSAS:90 and Puerto Rico was added as a self-representing PSU. Within NPSAS PSUs, institutions were selected with probabilities proportional to a measure of size related to enrollment. A total of 1,533 institutions were sampled. Systematic samples of students were drawn from these institutions at four sampling points: August 1, 1989; October 15, 1989; February 15, 1990; and June 15, 1990. A total of over 70,000 students were selected.

Response Rates and Their Derivations

Response rates were computed as the ratios of the number of sampled units that completed the survey over the number of eligible units in the sample. Ineligible units were deleted from the sample before data collection. Therefore, they were not included in the denominator in calculating response rates.

The overall weighted response rate for institutions was 86 percent. The rate was somewhat higher for certainty institutions at 91 percent than non-certainty institution at 86 percent. The overall weighted participation rate for students included in the analysis file was 84 percent (weighted number of students in the analysis file divided by the weighted number of eligible students); the overall weighted response rate for the student.CATI was 76 percent. The weighted response rate for students' parents to the parent CATI was 87 percent (weighted number of students' parents who participated divided by the weighted eligible number of students' parents).

1.3 Instrument Design

Descriptions of Instruments

Because of the complex nature of student finances, no single source of data is sufficient for studying financial aid. Higher education institutions are the best source for information on how they made financial aid decisions in determining a student's eligibility for aid and the amount of aid. Higher education institutions also can provide a record-keeping system that promises much higher accuracy in some areas than might be expected of students; for example, a student may have difficulty in recalling the exact amount of financial aid received or may not be fully informed of the source of the aid (e.g., whether student earnings were classified as work-study, and if so, what the source of the work-study funds was). Students are the best source of information about their costs of education (institutions often use general models, rather than specific information about individual students' costs), their financial resources, and about many of their personal characteristics and attitudes. Parental information is important especially when a student is dependent and unaided, because students and institutions often lack full information about the parents' finances and attitudes.

NPSAS was designed to include separate institutional, student, and parent components so that a complete picture of financial aid could be obtained. In general, each component focused primarily on collecting information that could not be collected as accurately from other sources. However, some redundancy was purposely built into the questionnaires. For example, though institutional records were considered the best source of data on financial aid awards, a lesser amount of award data was also collected from students. The student data thus provide a means for comparing the institutional data with student self-report data, and helping to complete gaps if (1) no institutional data were collected, (2) the student also attended additional schools for which no institutional records had been examined, or (3) the student obtained financial aid from outside sources (e.g., employers, family, or private organizations) about which the institution was uninformed. As another example, some types of data (such as a student's race or ethnicity) could not be obtained at sufficiently high response rates from only a single source (many institutions do not collect such data, while many students refuse to provide such data in a survey), but the combined responses from both sources helped to provide much more complete data. In the particular case of students' racial/ethnic categories, data were obtained from 75 percent of students through the student survey, but an additional 13 percent of students could be classified from institutional record data, and 5 percent from multiple sources.

Following are brief descriptions of the types of data collected through each instrument.

Records Data. A major purpose of the record abstract was to collect information on financial aid awards. Separate sections of the abstract collected data on federal, state, institutional, veteran's,

graduate, and other awards, in a total of more than 50 individual aid categories. A second purpose was to collect data on how institutions determined a student's eligibility for aid: this included both the original data submitted on a student's financial aid application (such as the Student Aid Report [SAR] or Graduate and Professional School Financial Aid Service [GAPSFAS]), and the components of financial aid formulas for aid (the Pell budget formula, the Congressional Methodology, and a separate institutional methodology for those few schools that use one) and their results (Student Aid Index [SAI] and Expected Family Contribution [EFC]). Finally, data were collected from institutional records on a student's enrollment status and some student characteristics.

Student Survey. Many types of data were collected through the student survey.

- First, information was collected about a student's enrollment status, including all schools and terms attended during the NPSAS year. Though some of this information duplicated that in the record abstract, the student survey provides the only source of data about changes in enrollment status from one term to another, and about students' attendance at other schools.

- Second, the student survey collected students' self-reports of their financial costs. Again, this information was unique to the student survey, although the record abstract collected information about tuition costs and institutions' budgets for estimating student costs.

- Third, detailed information was collected about a student's income. In the case of financial aid, the information was generally less detailed than that collected in the record abstract, but the student survey provides the only source of data about financial aid from other schools that were attended, and about some outside sources of financial aid not known to the institution (e.g., friends and relatives, and some private organizations). Other income data that were collected included employment income, savings, income of the spouse, and aid from parents. Information was also collected about a student's eligibility for aid; for aided students, these data duplicate application data collected in the record abstract, but the student survey generally provides the only source of such information for unaided students.

- Fourth, the student survey was the primary source of data about student characteristics, including race/ethnicity, citizenship, military status, disability, age, choice of school, remediation, religion, educational expectations (highest degree sought) and parental characteristics such as education, income, and number of dependents. Some information on student characteristics was collected through the record abstract as well, but more extensive information was collected through the student survey since students were considered the best source of reliable data.

- Finally, the student survey collected information about students' attitudes, including their values, self-perceptions, and reasons for refusing financial aid. This information was available from the student survey only.

Parent Survey. Three types of data were collected through the parent survey. First, parents were asked to describe what financial support they had provided to the sampled student, in terms of

the dollar amount provided, the form in which it was provided (e.g., contribution or loan), and the source of funds. Second, they were asked about other dependents that they provided support to, including the total number of dependents, the number that ever attended college, the total tuition paid, and the tuition paid for elementary or secondary school. Finally, they were asked to describe their own financial condition, in terms of their employment income, type of job, average monthly expenses, total assets, and the amount of money they had borrowed to provide financial aid to the student. Because parents were considered the best source of these three types of data, information collected on parents from the record abstract and student survey was much less extensive.

Records Update. The records update was designed to collect a limited amount of updated information from institutional records on financial aid awards. The records update was preprinted with limited data from the record abstract, with blanks for correcting or updating the final award amounts and award periods, and for entering other aid not noted on the form. The form also collected information on the dependency status of the student, on the terms of enrollment, and on Pell or Institutional budgets used to determine financial aid eligibility. A new item contained on the records update but not on the record abstract was the amount of federal College Work-Study aid that was earned, in addition to the already recorded amount that had been awarded.

1.4 Data Collection Methods

Record Abstract Survey

One major component of NPSAS was the collection of data on students from institutional records. Typically, the data were in more than one location at each institution, with data being obtained from both the registrar and financial aid office, and sometimes from other sources, such as individual departments for information on graduate students. Data were also kept in varying formats.

An institutional coordinator was designated at each institution, who provided information on an Institutional Checklist concerning the sources and locations of the requested data. Because of the diversity of records and formats used, a Record Abstract was developed to standardize the data collection. This abstract contained separate sections on financial aid awards (i.e., federal, state, institutional, veteran's, graduate, and other awards, using a total of more than 50 individual aid categories), the original data submitted on a student's financial aid application (such as the Student Aid Report [SAR] or Graduate and Professional School Financial Aid Service [GAPSFAS]), data on financial aid formulas used to assign aid (the Pell budget formula, the Congressional Methodology, and a separate institutional methodology for few schools that use one), the results of using those formulas (Student Aid Index [SAI] and Expected Family Contribution [EFC]), and data on a student's enrollment status and some student characteristics. Data on financial aid awards and on financial aid

budgets were collected for each separate term. The Record Abstract was also designed to collect locating information for both the student and his/her parents for the student and parent surveys.

Trained data collectors visited the institutions, using the Institutional Checklist and Record Abstract to collect the data. Data collection was conducted during two periods--December 1989 through February 1990, and August 1990 through November 1990--depending on the term being sampled.

Depending on the student and institution, varying amounts of information were available. Information on a student's enrollment status could be obtained for essentially all students, while the availability of some types of student characteristics (e.g., race/ethnicity) depended on the institution. Student financial aid application information was collected for about 26,400 students for whom the data were available. Financial aid award information was collected--either from the institution or the student telephone interview--on about 35,500 of the 68,929 eligible students.

Record Update Task. For those students for whom record abstracting was performed during the first data collection period (December 1989 - February 1990), a second institutional visit was made to update the information as of June 1990. This updating was performed during the same time period when institutions were visited for record abstracting of students enrolled in February or June 1990.

The records update was preprinted with limited data from the completed record abstract, with blanks for correcting or updating the final award amounts and award periods, and for entering other aid not noted on the form. The form also collected information on the (updated) dependency status of the student, on the terms of enrollment, and on Pell or Institutional budgets used to determine financial aid eligibility. A new item contained on the records update but not on the record abstract was the amount of federal College Work-Study aid that was earned; only the amount awarded had been collected previously.

Student Survey. The student survey was conducted by telephone. The use of computer-assisted telephone interviewing allowed on-line editing of the data as the questionnaire was being completed, and the use of complex skip patterns that were indiscernible to the students and interviewers. (For detail, see section 1.3.)

Parent Survey. Like the student survey, the parent survey was also conducted by telephone using computer-assisted telephone interviewing. The survey was completed with a response rate of 87 percent (weighted). The parent survey was directed primarily at parents of undergraduate, dependent, non-aided students. (For detail, see section 1.3.)

CHAPTER 2. INSTITUTION SAMPLING AND ENLISTMENT

This chapter discusses the methods used to sample and enlist institutions. The steps involved in sampling institutions include: identifying the areas of the country for drawing the institution sample, constructing a frame of eligible institutions, and selecting a probability sample of institutions within selected areas. The sampled institutions were contacted by mail and followed up by phone. Participating institutions received enlistment packages explaining the purpose of the survey and details for participation. In the following sections, we describe these processes and the overall participation rate.

2.1 Area Sampling

NPSAS PSUs

Area sampling was used to reduce data collection costs and to maintain comparability with the 1987 NPSAS sample. By consolidating field data collection efforts to a limited number of areas, we reduced the cost of data collection activities, such as record abstracting from financial aid offices. Also, the 1987 NPSAS PSUs were used again to facilitate comparisons between the two NPSAS surveys. The use of the same PSUs improved the precision for estimates of change between 1987 and 1990, and took advantage of the field experience and resources developed during 1987 NPSAS.

The first stage of sampling was the selection of areas of the country, called primary sampling units (PSUs). A PSU was defined as a collection of geographically contiguous three-digit ZIP code areas. The 50 states, the District of Columbia, and Puerto Rico formed a total of 362 PSUs, none of which crossed state boundaries. The PSU sample originally selected for 1987 NPSAS consisted of 120 of these PSUs and was used again for NPSAS:90 (see Appendix E for additional information from the *Methodology Report for the National Postsecondary Student Aid Study, 1987* [NCES 90-309]). In addition, Puerto Rico was added as a self-representing PSU totaling 121 PSUs in NPSAS:90. Of these 121 PSUs, 51 were self-representing PSUs (selected with probability equal to unity), and 70 were non-self-representing PSUs (selected with probability less than unity).

Addition of Puerto Rico

The inclusion of Puerto Rico as a self-representing PSU was consistent with the methods used for defining self-representing PSUs in NPSAS:87 because Puerto Rico housed a large number of postsecondary institutions and students. In comparison with other PSUs, Puerto Rico was the third largest in terms of the number of institutions in the area, and the fourth largest in the number of

postsecondary students. Puerto Rico also had about seven times as many institutions as the average PSU, and six times as many students.

2.2 Institution Sample

Institution Eligibility

The second stage of the sample design was the sampling of institutions within the NPSAS PSUs. To be eligible for NPSAS:90, an institution satisfied all the conditions listed in figure 2.2.1. The main exclusions were institutions that provided only avocational, recreational, or remedial courses; offered only in-house business courses; offered only seminars of less than three months duration (such as driver training schools, real estate courses, and tax preparation classes); or offered only correspondence courses. United States Service academies were not eligible because of their unique funding and tuition situation. Among the 1,533 institutions selected for NPSAS:90, 285 (about 20 percent) were ineligible and therefore excluded from the survey.

Institution Frame Building

The sampling frame for institutions was based primarily on the 1987-88 IPEDS Institution Characteristics (IC) file. The list of institutions on this file was checked against a number of other sources to improve coverage, to eliminate ineligible institutions, and to collapse multiple listings. To improve coverage, institutions that were not in the IC file, but appeared in: (a) the 1987 Fall Enrollment (IPEDS-EF) file, (b) the 1987-88 file of Pell Grant participating institutions, or (c) were identified as eligible for NPSAS during initial contacts with NPSAS institutions, were added to the frame.

Some ineligible institutions were deleted from the frame before institutions were sampled. The type of institutions eliminated were U.S. service academies, central offices or administrative buildings, nonexistent or closed institutions, and classrooms misidentified as institutions. Other ineligible institutions such as schools offering programs like real estate schools and correspondence schools remained in the frame. The IPEDS data base did not provide detailed information about the specific length of programs less than 600 hours, thus many proprietary institutions had to be contacted before the programs they offered could be identified.

Lastly, some institutions that offered a variety of programs might be listed several times in the frame. For example, a hospital offering training in radiologic technology, nursing, and emergency medical technician training, might be listed three times. To improve efficiency in sampling and field

operations, institutions offering more than one program at the same address were collapsed into one entry.

Table 2.2.1 shows the number of institutions added and subtracted from the IPEDS-IC file during the process of frame development. There were 12,243 institutions in the original IPEDS-IC file. We added 745 institutions from other sources, eliminated 58 ineligible institutions, and deducted 62 multiple entries. Thus, the final NPSAS:90 institutional frame consisted of 12,868 institutions.

Figure 2.2.1. Institutions Eligible for NPSAS:90

Institutions in NPSAS:90 satisfied all of the following conditions at the beginning of the 1989-90 school year:

- Offered an education program designed for persons who have completed secondary education;
- Offered an academically, occupationally or vocationally oriented program of study;
- Offered access to persons other than those employed by the institution;
- Offered more than just correspondence courses;
- Offered at least one program lasting three months or longer; and
- Were located in the 50 states, Puerto Rico or the District of Columbia.

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

Table 2.2.1. Frame Development: Number of institutions in the 1987-88 IPEDS-IC File and in the NPSAS:90 institutional frame

Educational Sector	Number of Institutions:				
	In IPEDS-IC File	Added to IC File[a/]	Eliminated from IC File[b/]	Collapsed Multiple Entries	In NPSAS:90 Institutional Frame
Public, 4-year and above	638	9	12	2	633
Private, nonprofit, 4-year or above	1,944	49	3	7	1,983
Private, for-profit, 4-year or above	120	7	1	1	125
Public, 2-year	1,257	81	4	1	1,333
Private, nonprofit, 2-year	845	88	4	35	894
Private, for-profit, 2-year	850	49	8	0	891
Public, less-than-2-year	380	52	1	2	429
Private, nonprofit, less-than-2-year	515	43	1	14	543
Private, for-profit, less-than-2-year	5,694	367	24	0	6,037
Total	12,243	745	58	62	12,868

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

a/ Institutions that were not in the 1987-88 IPEDS-IC file but listed in the 1987 IPEDS-Fall Enrollment file, or the 1987-88 Pell Grant institution file were added.

b/ Ineligible institutions such as central offices, nonexistent or closed buildings, and U.S. service academies were eliminated.

2.3 Institution Sample Selection

Sampling Procedure

The sampling procedure within NPSAS PSUs involved classifying institutions into strata, allocating a sample size per stratum, assigning measure of size (MOS) to each institution, and selecting a systematic sample of institutions with probabilities proportionate to the MOS from each stratum across the PSUs. A hierarchical system of 14 strata were defined depending on the type and control of institutions, and by the number of first-professional, doctoral, master's or baccalaureate level students enrolled in the highest degree programs offered by the institutions. Table 2.3.1 shows the definition of the sampling strata and the number of institutions sampled per stratum. Of the total of 1,533 institutions in the sample, 339 were certainty institutions (selected with probability equal to unity) and 1,194 were non-certainty institutions (selected with probability less than unity).

The certainty institutions were self-representing institutions selected from the entire frame and from self-representing NPSAS PSUs. Table 2.3.2 shows the number of self-representing and non-self-representing institutions selected from the entire frame, self-representing NPSAS PSUs, and non-self-representing NPSAS PSUs. Among the 339 certainty institutions, 225 were selected from the entire frame and 114 from self-representing NPSAS PSUs. Other self-representing institutions from non-self-representing NPSAS PSUs were not certainty institutions because the probabilities of selecting the PSUs were less than unity. The 1,194 non-certainty institutions consisted of 502 non-self-representing institutions from self-representing NPSAS PSUs, 250 self-representing institutions, and 442 non-self-representing institutions from non-self-representing PSUs.

Certainty institutions selected from the entire frame were institutions with a total enrollment above a prescribed cutoff per sampling stratum. These institutions were included to ensure that large institutions in the nation were represented in the survey. The sampling intervals for the other institutions in NPSAS PSUs were determined after the large certainty institutions were removed from the frame. Institutions with enrollment higher than the sampling interval for the stratum were sampled as self-representing institutions, and this was done before the sampling of non-self-representing institutions. The enrollment cutoff for selecting self-representing institutions from the frame and the sampling intervals for institutions in NPSAS PSUs are shown in table 2.3.3.

For non-certainty institutions, the probability of selection p_{jk}, was equal to:

$$P_{jk} = \frac{n_j \, m_{jk}}{\sum_k m_{jk}}$$

where n_j was the number of institutions in stratum k, and m_{jk} was the MOS for institution j in stratum k. The probability for the first stage selection of PSUs is not shown in this equation for clarity of presentation.

The measure of size, m_{jk}, was a composite measure derived from the number of students enrolled in first-professional, graduate, and undergraduate programs. This measure was the weighted sum of the number of students in each educational level. The weights used for first-professional, graduate, and undergraduates were five, two, and one respectively. Institutions with a small composite MOS were assigned a minimum value which was determined by examining the distribution of enrollment totals per stratum. This composite measure gave approximately self-weighting samples for multiple domains in multi-stage samples. The advantages were: it ensured adequate coverage for each study domain, it controlled for the precision of estimates; and it was flexible with regard to changes in sample sizes, sampling rates, and population counts.

The enrollment figures used to compute the measure of size were taken from the 1987-88 IPEDS IC and EF files. A substantial amount of effort was devoted to checking the consistency of these files, and to update the enrollment values based upon other published reports (such as the HEP) or upon telephone contacts with institutions[1]. In addition to the primary sample of institutions, it was decided that a supplemental sample of institutions participating in the Stafford Loan programs (GSL) was necessary to ensure coverage of small institutions. The 1987-88 file of institutions participating in the Stafford Loan programs consisted of 16,385 records; including 7,598 institutions located in NPSAS:90 PSUs. Because the Stafford file did not provide any measure of enrollment, the sampling rate established for less-than-2-year private for-profit institutions in the main sample were used to draw an equal probability sample of supplemental institutions. The sampling rate used was approximately 1 in 30. This sampling scheme resulted in an initial selection of 476 institutions which were checked against the listing of institutions in the NPSAS survey frame. Any supplemental institutions that were found in the NPSAS frame were excluded to avoid the problem that an institution might have two chances of being selected. Institutions defined as being closed were also excluded from the sample at this time. This process reduced the supplemental sample to 81 institutions. Of the 81 institutions dispatched to the field for clarification and unduplication only 9 were eligible for NPSAS:90.

---

[1]Enrollment counts were available from the 1987-88 IPEDS IC and EF files for over 95 percent of the institutions in the frame. However, the availability of data varied substantially by school types and control. In the sector of less-than two-year institutions, the percentage of public institutions with enrollment data was 96 percent, the percentage for private, not-for-profit schools was substantially lower at 57 percent, and that for private, for-profit schools was 60 percent.

Table 2.3.1. Institution sample: Number of institutions in the frame and in the sample

Sampling Stratum	Number of Institutions in:			
	Frame	Sample		
		Certainty	Non-certainty	Total
Public, 4-year, first-professional[a/]	140	55	27	82
Private, 4-year, first-professional[a/]	279	56	46	102
Public, 4-year, doctor's degree[b/]	95	25	19	44
Private, 4-year, doctor's degree[b/]	120	18	18	36
Public, 4-year, master's degree[c/]	245	49	43	92
Private, 4-year, master's degree[c/]	349	13	61	74
Public, 4-year, bachelor's degree[d/]	155	6	23	29
Private, 4-year, bachelor's degree[d/]	1,362	3	86	89
Public, 2-year	1,327	23	188	211
Private, 2-year, not-for-profit	887	16	84	100
Private, 2-year, for-profit	891	12	98	110
Public, less-than-2-year	429	4	78	82
Private, less-than-2-year, not-for-profit	552	22	83	105
Private, less-than-2-year, for-profit	6,037	37	340	377
Total	12,868	339	1,194	1,533
Supplemental institutions[e/]	16,385	0	81	81

a/ These institutions were schools that offered first-professional programs and had a first-professional enrollment greater than 75 in 1987.

b/ These institutions were schools that offered programs up to doctorate level, had a combined first-professional and graduate enrollment greater than 75, and had failed one or more of the inclusion criteria for the above strata.

c/ These institutions were institutions that offered degrees or certificates up to post-baccalaureate or post-master's level, had a combined first-professional and graduate enrollment greater than 75, and had failed one or more of the inclusion criteria for the above strata.

d/ These institutions were institutions that offered programs up to the bachelor's level, and had failed one or more of the inclusion criteria for the above strata.

e/ Supplemental institutions were selected from the 1987-88 institution file for Stafford loan programs.

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

Table 2.3.2.--Institution sample: Number of self-representing and non-self-representing institutions from the entire frame, self-representing NPSAS PSUs and non-self-representing NPSAS PSUs

Sampling Stratum	Entire Frame	Self-Representing PSUs		Non-Self-Representing PSUs	
		Number of Institutions:		Number of Institutions:	
	Number of Self-Representing Institutions	Self-Representing	Non-Self-Representing	Self-Representing	Non-Self-Representing
Public, 4-year, first professional[a/]	38	17	10	16	1
Private,4-year, first professional[a/]	42	14	27	13	6
Public, 4-year, doctor's degree[b/]	6	19	0	19	0
Private, 4-year, doctor's degree[b/]	4	14	12	5	1
Public, 4-year, master's degree[c/]	22	27	8	32	3
Private, 4-year, master's degree[c/]	12	1	31	18	12
Public, 4-year, bachelor's degree[d/]	3	3	9	10	4
Private, 4-year, bachelor's degree[d/]	3	0	28	14	44
Public, 2-year	13	10	68	45	75
Private, 2-year, nonprofit	14	2	32	11	41
Private, 2-year, for-profit	12	0	43	14	41
Public, less-than-2-year	4	0	15	31	32
Private, less-than-2-year, nonprofit	15	7	45	10	28
Private, less-than-2-year, for-profit	37	0	174	12	154
Total	225	114	502	250	442

a/ These institutions were schools that offered first-professional programs and had a first-professional enrollment greater than 75 in 1987.

b/ These institutions were schools that offered programs up to doctorate level, had a combined first professional and graduate enrollment greater than 75, and had failed one or more of the inclusion criteria for the above strata.

c/ These institutions were institutions that offered degrees or certificates up to post-baccalaureate or post-master's level, had a combined first-professional and graduate enrollment greater than 75, and had failed one or more the inclusion criteria for the above strata.

d/ These institutions were institutions that offered programs up to the bachelor's level, and had failed one or more of the inclusion criteria for the above strata.

e/ Supplemental institutions were selected from the 1987-88 institution file for Stafford loan programs.

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

BEST COPY AVAILABLE

Table 2.3.3. Institution sample: Enrollment cutoff for self-representing institutions from entire frame and sampling intervals for institutions in NPSAS PSUs

Sampling Stratum	Enrollment Cutoff for Self-Representing Institutions from Entire Frame	Sampling Intervals for Institutions in NPSAS PSUs
Public, 4-year, first-professional[a/]	24,000	10,742
Private,4-year, first-professional[a/]	9,000	13,201
Public, 4-year, doctor's degree[b/]	24,000	--
Private, 4-year, doctor's degree[b/]	9,000	7,873
Public, 4-year, master's degree[c/]	14,000	9,634
Private, 4-year, master's degree[c/]	6,000	10,188
Public, 4-year, bachelor's degree[d/]	14,000	9,992
Private, 4-year, bachelor's degree[d/]	6,000	9,879
Public, 2-year	22,000	21,746
Private, 2-year, nonprofit	1,500	1,357
Private, 2-year, for-profit	1,900	2,569
Public, less-than-2-year	2,200	1,406
Private, less-than-2-year, nonprofit	500	422
Private, less-than-2-year, for-profit	1,400	1,531

a/ These institutions were schools that offered first-professional programs and had a first-professional enrollment greater than 75 in 1987.
b/ These institutions were schools that offered programs up to doctorate level, had a combined first-professional and graduate enrollment greater than 75, and had failed one or more of the inclusion criteria for the above strata.
c/ These institutions were institutions that offered degrees or certificates up to post-baccalaureate or post-master's level, had a combined first-professional and graduate enrollment greater than 75, and had failed one or more the inclusion criteria for the above strata.
d/ These institutions were institutions that offered programs up to the bachelor's level, and had failed one or more of the inclusion criteria for the above strata.
e/ Supplemental institutions were selected from the 1987-88 institution file for Stafford loan programs.

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

2.4 Institution Enlistment

Westat informed the Chief Administrator at the sampled institutions that their institution had been selected and requested participation in the project. Non-response follow-up was done by telephone with additional follow-up by phone for refusal conversion of those schools that initially declined participation.

2.4.1 Advance Packet Mailing

Introductory packages of NPSAS information were sent to the Chief Administrators of all sampled institutions beginning June 8, 1989 and continuing on a flow basis over the next several weeks. The mailout informed institutions of their selection for the study. Several other goals of this initial mailing are described below.

A cover letter from NCES with the signature of Emerson Elliott, the Acting Commissioner, described the study and its sample and requested the school's participation. It introduced Westat as the contractor for the study and requested that the Chief Administrator appoint a coordinator to the study and return to Westat an enclosed postcard on which the name of the coordinator was to be written. The letter also provided the names and phone numbers of the NCES co-project officers and the Westat director of survey operations if the institution coordinator had any questions about the study.

In addition to this letter, materials were included which described details of the data collection process, estimates of time commitments required by the study of the institution, a statement of authorization and confidentiality, a summary brochure of the study, and a business reply postcard. The postcard requested verification of the school's name and address, the name and title of the appointed coordinator, identification of control (public; private, nonprofit; private, for-profit) and highest degree offered, and an estimate of 1989-90 undergraduate enrollment. The last three items provided verification of information needed in order to select the sampling rates for each school. An example of the advance packet is included in the Appendix.

2.4.2 Institution Participation Receipt Control System

The receipt system used to track the mailout and monitor the institutions' participation throughout the various components of the study was similar to the system used in the 1987 study. The system included preloaded information about the sampled institutions (e.g., institution name and Westat assigned ID; participation status in the 1987 NPSAS, if relevant; level and control as reported to IPEDS; sampling points; and several other pertinent variables). The system also provided blank fields to report on an institution's 1990 participation status, any updates by the institution of the pre-loaded variables, and many other institution-specific fields (coordinator name and institution sample size, for example) which were used throughout the study.

Upon receipt of the postage-paid postcards, operations staff entered both the date of receipt and the information provided on the cards into the receipt system. Initial participation statuses (participating or initial refusals) were assigned at this point. While most of the returned postcards were filled in correctly, there were some problems. The most common of these were institutions not understanding the terminology (institutions whose control should have been identified as "private, for-profit" indicating they were a nonprofit institution because they had, in fact, never earned a profit); institutions misreading the item cues ("< 2-year only" read without the less-than sign, as "2-year only"); and lack of specificity in the item request (exactly what time frame was being referred to when requesting an estimate of 1989-90 undergraduate enrollment). These problems were resolved during further contacts with the institution.

If postmaster returns included a forwarding address or an address correction or operations staff was able to determine a valid address through directory assistance or other sources, staff assigned a re-mail status. A new address label was then generated, the package was re-mailed and the date of the re-mail was entered. Postmaster returns for which no new address information could be obtained were assigned a final non-participating status of non-locatable.

2.4.3 Institution Nonresponse Follow-up

A second mailing of enlistment materials and a brief reminder letter was prepared and sent to about 900 nonresponding institutions. This mailing began on June 30, 1989 and continued through the first week of July.

On July 24, telephone follow-up began at Westat's Telephone Research Center (TRC) for nonresponding institutions. Nine interviewers and one supervisor were trained in using the non-response follow-up script to enlist institutions and obtain the requested information. The four-hour training session consisted of an overview of the project, a review of purposes and procedures for the nonresponse follow-up, and administrative procedures. Role-play activities were performed with the trainers playing the part of institution staff and the interviewers using the script to enlist institutions and obtain the requested information. Interviewer calls to institutions were monitored randomly over the following week with small group discussions held, as necessary, to clarify procedures.

All calls were documented using standard TRC procedures. Problem cases were also documented, reviewed by the supervisor and, if necessary, given to operations staff to resolve. All cases receiving a finalized status from the TRC, whether participating or non-participating, were returned to the operations area where updated statuses and information from the completed script were entered into the receipt control system.

returned to the operations area where updated statuses and information from the completed script were entered into the receipt control system.

Of the participating schools, approximately 71 percent were successfully enlisted by mail. (That is, the return postcard was received with the requested information, including the name of the coordinator.) Successful enlistment of institutions and identification of a NPSAS coordinator for the remainder of the participating schools was completed by telephone followup. Table 2.4.1 displays the source of institution enlistment by level and control.

2.4.4 Institution Coordinator Packet Mailing

As institutions were assigned a participating status and institution coordinators for the study were identified packets of informational materials were mailed to coordinators. This mailing began on September 8 and continued over the next month as additional institutions agreed to participate.

The coordinator packet included a cover letter to the coordinator, a copy of the packet of materials sent to the Chief Administrator of the institution, an envelope containing student enrollment list labels and business reply labels, an institution-specific, computer-generated institution background data verification (IBDV) sheet, enrollment list request information (variable depending on the institution's level and the number of times it was to be sampled), a tape/diskette transmittal sheet (for those schools wishing to send enrollment lists using this format), and a NPSAS folder used to hold all of these materials.

Coordinators were given the telephone number for the NPSAS Information Line (an 800 number) and asked to call if any of the information contained on the institution background data verification sheet was incorrect, if there were any problems related to the request for enrollment lists, or if they had any questions about the study.

Beginning on September 25, those institutions which had been mailed a coordinator packet but had not yet called in to Westat were contacted by the Telephone Research Center to verify receipt of the package. Additionally, this call was used to verify the information printed on the IBDV, to review enrollment list requests, and to prompt for August enrollment lists, if relevant. These calls continued through October.

At participating institutions, approximately 51 percent of the persons appointed as NPSAS coordinators were staff within the institution's financial aid office. Others appointed coordinators were divided fairly evenly among CEOs/presidents/owners and the offices of admissions, the

registrar, student affairs and/or services, and institutional research and planning. About 75 percent of the participating institutions remained with the same coordinator through all phases of the study.

2.4.5 Refusal Conversion

Refusal conversion procedures for institutions assigned a status of initial refusal (IR) began in September 1989 and, because schools could refuse to participate at any point during the study, continued throughout the duration of the study. The bulk of the refusals were obtained during the enlistment phase but additional refusals occurred during initial coordinator contact, scheduling of field visits, requests for additional sampling lists, and contact and scheduling for the Records Update Task.

The approach for handling refusal conversions during the enlistment phase of the study was three-pronged, involving Telephone Research Center (TRC) staff trained for refusal conversion, Westat senior operations staff, and NCES. Institutions which were assigned an "IR" status because they returned an incomplete postcard with a note saying that they did not participate in the Department of Education's financial aid programs and nonresponding institutions which were suspected of being ineligible for NPSAS were given to trained TRC staff to initiate conversion and determine eligibility. Institutions coded "IR" as a result of a letter to NCES or Emerson Elliott, Acting Commissioner, in which they declined to participate or as a result of an 800 number phone call in which the Chief Administrator or his appointee spoke with senior-level operations staff and declined, were collated and reviewed by operations staff. Depending on the reason for the refusal and the potential impact on the study if accepted, schools were either given a final refusal status, turned over to NCES for refusal conversion or finalization of the refusal status, or divided among senior level operations staff for refusal conversion. Senior operations staff did the refusal conversion.

2.5 Institution Participation Rates

The number of institutions that participated in NPSAS:90 are shown in table 2.5.1. The weighted and unweighted response rates for institutions were computed as the ratio of the number of institutions that completed the survey over the number of eligible institutions in the sample. Chapter 9 describes weighted and variance estimation procedures. Institutions that were regarded as ineligible during the initial screening were not included in the denominator.

The overall unweighted response rate of all institutions was 91 percent, and the weighted response rate was slightly lower at 86 percent. In general, the response rates for public institutions were higher than those of private institutions. The unweighted response rates for public institutions ranged between 92 and 95 percent, those for private institutions ranged between 85 and 92 percent.

The weighted response rates for these school types were more variable, ranging from 89 percent to 96 percent for private institutions, and 80 to 93 percent for public institutions. The sector with the lowest response rate was private for-profit institutions, the response rate of which tended to be less than 90 percent.

While unweighted response rates for certainty and non-certainty institutions were similar, the weighted response rate for certainty institutions at 91 percent was higher than that of non-certainty institutions at 86 percent.

Table 2.4.1. Source of Institution Enlistment Status for Participating Institutions

Control and Level of Institution	Source					
	Returned Postcard		Telephone Followup		Total	
	Number of Institutions	Percent of Institutions	Number of Institutions	Percent of Institutions	Number of Institutions	Percent of Institutions
Total	798	70.6	332	29.4	1,130	100.0
Public						
Total	383	33.9	96	8.5	479	42.4
4-year doctoral	97	8.6	18	1.6	115	10.2
Other 4-year	100	8.8	13	1.2	113	10.0
2-year	141	12.5	48	4.2	189	16.7
Less-than-2-year	45	4.0	17	1.5	62	5.5
Private, non-profit						
Total	273	24.2	94	8.3	367	32.4
4-year doctoral	92	8.1	36	3.2	128	11.3
Other 4-year	114	10.1	32	2.8	146	12.9
2-year	47	4.2	12	1.1	59	5.2
Less-than-2-year	20	1.8	14	1.2	34	3.0
Private, for-profit						
Total	142	12.6	142	12.6	284	25.2
4-year doctoral	0	0	2	.2	2	.2
Other 4-year	2	.2	2	.2	4	.4
2-year	42	3.7	30	2.7	72	6.4
Less-than-2 year	98	8.7	108	9.6	206	18.2

Table 2.5.1. Institutional response rates: Number and percentage of institutions participating in NPSAS:90

Institutional Sector	Number of:		Response Rates	
	Respondents	Non-Respondents	Unweighted	Weighted
Public, less-than-2-year	62	3	95%	89%
Public, 2-year	189	16	92%	96%
Public, other 4-year	113	7	94%	92%
Public, doctoral	115	9	93%	93%
Private, not-for-profit, less than 2-year	34	3	92%	93%
Private, not-for-profit, 2-year	59	10	86%	86%
Private, not-for-profit, other 4-year	146	13	92%	88%
Private, not-for-profit, doctoral	128	12	91%	91%
Private, for-profit, less-than-2-year	206	37	85%	80%
Private, for-profit, 2-year or more	78	8	91%	87%
Certainty institutions	259	26	91%	91%
Non-certainty institutions	871	92	90%	86%
Total	1,130	118	91%	86%

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

CHAPTER 3. STUDENT SAMPLING

This chapter discusses the selection of students within sampled institutions. In order to obtain a full-year sample of students, sampling was done in the fall (October 1989) and in three nonfall periods (August 1989, February 1990, and June 1990). Students were selected from all NPSAS:90 institutions during the fall sampling, but only a subsample of institutions were involved with the nonfall sampling. The detailed sampling procedures and quality control processes are discussed below.

3.1 Student Sampling

Student Eligibility

The first two stages of sampling in NPSAS:90 involved institutions and are discussed in Chapter 2. The third stage of sampling was the selection of students within sampled institutions. To be eligible for NPSAS:90, a student must have attended an eligible institution during the 1989-90 school year and been enrolled in one or more of the following: course(s) for credit; degree or formal award program of at least three months duration; or occupationally or vocationally specific program of at least three months duration. Students enrolled in a high school program were excluded regardless of whether they satisfied the above conditions. The eligibility criteria for the student sample are summarized in figure 3.1.1. These eligibility criteria encompassed nearly all postsecondary students. The main exclusions were students in correspondence courses or in programs of very short duration.

Sampling Procedure

The basic design for the sampling of students involved the following steps: collect lists of all students enrolled in the academic year 1989-90 from the sampled schools, stratify students by educational level (undergraduate, graduate, and first-professional) and by school type and control (10 strata), determine sampling rate per stratum to achieve the desired sample size, and apply rate to select a systematic sample of students. Most lists were sorted by student last name in alphabetical order. The expected sample size per stratum to achieve a total sample of about 70,000 students is shown in table 3.1.1.

The actual implementation of this plan, however, was complicated by two issues. First, while institutions could provide lists of students enrolled in a particular term, they were unable to provide unduplicated lists of students enrolled in more than one term. In order to obtain a full-year sample of students, sampling was done based on enrollment as of the following: August 1, 1989; October 15, 1989; February 15, 1990; and June 15, 1990. Second, this sample design required the co-operation

of institutions in providing the number of students enrolled and separate lists of students enrolled in undergraduate, graduate, or first-professional programs. Quality control procedures were established to ensure that information from institutions was updated and verified. For students who transferred to other institutions during the year, procedures were established for their correct classification. The following sections of this chapter expand on these issues and discuss the procedures for drawing the samples and quality control.

Figure 3.1.1. Students eligible for NPSAS:90

Postsecondary students who had attended a NPSAS eligible institution between
July 1, 1989 and June 30, 1990, and enrolled in one or more of the following courses or programs:

- course(s) for credit;
- degree or formal award program of at least three months duration; or
- occupationally or vocationally specific program of at least three months duration.
- Regardless of the above attendance status, if a student was also enrolled in a high school program, the student was excluded.

Table 3.1.1. Expected student sample for NPSAS:90: Sample size for the full-year by institutional sector and student educational level

Institutional Sector	Number of Students:			
	Total	Undergraduate	Graduate	First-Professional
Public, doctoral	12,886	7,036	3,650	2,200
Private, doctoral	15,933	9,543	2,590	3,800
Public, other 4-year	8,791	6,911	1,880	
Private, other 4-year	9,590	7,710	1,880	
Public, 2-year	7,530	7,530		
Private, not-for-profit, 2-year	2,259	2,259		
Private, for-profit, 2-year	2,710	2,710		
Public, less-than-2-year	1,800	1,800		
Private, not-for-profit, less-than-2-year	1,500	1,500		
Private, for-profit, less-than-2-year	7,201	7,201		
Total Sampled	70,200	54,200	10,000	6,000

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

3.2 Selection of the Fall and Nonfall Samples of Students

The Fall Sample

The fall sample of students was selected from October 15, 1989 enrollment lists. The students from the enrollment lists were first stratified by level of education (first-professional, graduate, and undergraduate) and by the level and control of the institution they attended. The number of students sampled per strata for the fall is shown in table 3.2.1. More than 60,000 students were sampled; of which, approximately 46,000 were undergraduates, 8,700 were graduates, and 5,600 were first-professional students.

The sampling rate used to select the fall sample of students was the ratio of the desired sample size divided by the estimated number of students in the population per stratum. In order to calculate this rate, the population size was initially estimated using information about the fall enrollment from the IPEDS IC and EF files. These estimates were updated and revised as contacts were made with institutions and the final sampling rates were computed after the adjustments. It was estimated that about 13 million students were enrolled in the fall; the estimated population size per sampling stratum is shown in table 3.2.2.

Since participating institutions were not selected with equal probabilities, the within-institution student sampling rate was the overall sampling rate described above divided by the probability of selecting the institution. Mathematically, R_{ij}, the within-institution sampling rate for students in institution j stratum i is,

$$R_{ij} = r_i / P_j,$$

where r_i is the overall ratio of sample size to population size in stratum i, and P_j is the probability of selecting institution j in the early stages of sampling. The within-institution sampling rate for small institutions was adjusted so that the minimum number of students sampled per institution was approximately 12 for the fall and 10 for nonfall. The sampling interval used to draw the systematic sample, I_{ij}, is $(1/R_{ij})$.

A large proportion of students in NPSAS:90 were sampled in the fall to ensure comparability with NPSAS:87 which was conducted on a sample of fall students only. The large fall sample in NPSAS:90 helped to improve the precision of estimates of change between the two surveys. The trade-off was that the precision for estimates for the full-year could be diminished. Because reliable estimates of the number of students not enrolled in the fall were not available it was impossible to fine tune the sampling rates to obtain optimal rates for fall and nonfall time periods.

The Nonfall Sample

To ensure coverage of the nonfall periods, additional samples of students were selected on August 1, 1989; February 15, 1990; and June 15, 1990. The basic sampling scheme used for the nonfall samples was similar to that of the fall sample. A total of about 14,000 students were planned to be sampled and the expected number of sampled students per stratum is shown in table 3.2.3.

To minimize the administrative burden of institutions, not all institutions that participated in the fall were used for nonfall sampling. Instead, institutions were separated into subgroups and different groups were sampled at different times. Table 3.2.4 shows the time periods in which different subgroups were sampled. Institutions offering 4-year or 2-year programs were divided into two equal groups. Group A institutions were sampled at all four sampling times. Group B institutions were used for fall sampling only. Less-than-2-year institutions were subdivided into 4 groups so that more institutions could be involved with nonfall sampling but fewer were used at all four sampling points[2]. About 40 percent of the less-than-2-year institutions were classified as group A and these institutions were sampled at all four sampling points. The remaining institutions were subdivided into three equal groups. Group B institutions were sampled in October, February, and June. Group C institutions were sampled in August and October. Group D institutions were sampled in October only.

Like the fall sample, the nonfall students were selected using a stratified systematic sampling scheme. The overall sampling rate for students in each student stratum was the ratio of the desired number of students in the sample over the number of students in the population. An adjustment factor, B_{ij}, was included to account for the subsampling of institutions in the nonfall periods. Mathematically, the within-institution nonfall sampling rate for students in student stratum i and institution j, R'_{ij}, is

$$R'_{ij} = r'_i / (P_j B_{jt})$$

where r'_i is the overall nonfall sampling rate for students in stratum i; P_j is the probability of selecting institution j, and B_{jt} is the probability of including institution j at time t. A minimum sample size of 10 students was imposed for the nonfall sample.

Although sampling was done at four points in time, students were allowed only one chance of being selected from each institution. To accomplish this, students sampled during the nonfall periods were checked against the entire fall sampling list within the institution. Sampled nonfall students who were on the fall sampling lists for the same institution were eliminated. Specifically, students

---

[2]From the experience with NPSAS:87, almost half of the students from less than 2-year institutions were not enrolled in the fall. Therefore, a larger sample of less then two-year institutions was needed for the non-fall samples.

sampled from the August list were checked against the entire sampling list for October. Those found in the October list were deleted from the August sample. Likewise, students sampled in February were checked against the October and the August lists; those sampled in June were checked against the August, October, and February lists. Students found in any prior sampling lists were deleted. This checking was done across levels, so that, students who changed education level during the year would still have only one chance of being selected from the institution. Students enrolled at another school during a non-sampled term were included in the study. For example, students selected from the fall term at institution X could have been selected from institution Y in August, February, or June.

The size of the June sample was reduced during actual data collection because of cost considerations. Only less-than-2-year institutions were used to draw the student sample and the sample size was reduced from about 200 less-than-2-year institutions to about 100 institutions, and from about 2,000 students to 500 students. The reduction of the June sample introduced a slight downward bias in estimates for students in 4-year and 2-year institutions, and increased the variability in estimates of students in less-than-2-year schools. However, as meaningful an adjustment as possible (based on the best information available) was made to correct this, based on the number of Pell grant recipients.

Table 3.2.1. Student sample for NPSAS:90: Sample size for the fall sample by institutional sector and student educational level

Institutional Sector	Number of Students:			
	Total	Undergraduate	Graduate	First-Professional
Public, doctoral	12,235	7,030	3,165	2,040
Private, doctoral	15,145	9,330	2,245	3,570
Public, other 4-year	8,400	6,770	1,630	
Private, other 4-year	9,165	7,535	1,630	
Public, 2-year	5,580	5,580		
Private, not-for-profit, 2-year	1,675	1,675		
Private, for-profit, 2-year	2,010	2,010		
Public, less-than-2-year	1,055	1,055		
Private, not-for-profit, less-than-2-year	880	880		
Private, for-profit, less-than-2-year	4,225	4,225		
Total Sampled	60,370	46,090	8,670	5,610

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

Table 3.2.2. Student sample for NPSAS:90: Estimated population size for fall enrollment by institutional sector and student educational level

Institution Level and Control	Number of Students in Thousands			
	Total	Undergraduate	Graduate	First-Professional
Public, doctoral	3,300	2,550	650	100
Private, doctoral	1,195	760	300	135
Public, other 4-year	1,850	1,630	220	
Private, other 4-year	1,320	1,200	120	
Public, 2-year	4,075	4,075		
Private, not-for-profit, 2-year	110	110		
Private, for-profit, 2-year	200	200		
Public, less-than-2-year	200	200		
Private, not-for-profit, less-than-2-year	40	40		
Private, for-profit, less-than-2-year	400	400		
Total	12,690	11,165	1,290	235

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

Table 3.2.3. Student sample for NPSAS:90: Expected number of students for the nonfall sample by institutional sector and student educational level

Institutional Sector	Number of Students:			
	Total	Undergraduate	Graduate	First-Professional
Public, doctoral	1,125	360	560	205
Private, doctoral	1,400	695	400	305
Public, other 4-year	775	490	285	
Private, other 4-year	845	560	285	
Public, 2-year	2,625	2,625		
Private, nonprofit, 2-year	790	790		
Private, for-profit, 2-year	945	945		
Public, less-than-2-year	960	960		
Private, nonprofit, less-than-2-year	810	810		
Private, for-profit, less-than-2-year	3,840	3,840		
Total	14,115	12,075	1,530	510

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

Table 3.2.4. Student sample for NPSAS:90: Sampling time points for groups of institutions

Level of Institution	Expected Number of Participating Institutions	Sampling Time Point			
		August	October	February	June
4-year institutions					
Group A	245	X	X	X	X[a/]
Group B	244		X		
2-year institutions					
Group A	177	X	X	X	X[a/]
Group B	176		X		
Less-than-2-year institutions					
Group A	149	X	X	X	X
Group B	69		X	X	X
Group C	70	X	X		
Group D	70		X		

a/ These institutions were not included in the actual June sample because of cost considerations.

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study.

3.3 Quality Control and Problem Handling in the Sampling Process

The sampling procedures described above depended upon knowledge about the institutions' level and control, enrollment sizes at each student level (undergraduate, graduate, and first-professional), and the cooperation from institutions in supplying lists of students enrolled for different levels of studies during the four sampling periods. To ensure that the sampling specifications were strictly followed, quality control procedures were established to update information and to resolve problems associated with sampling from enrollment lists supplied by institutions.

Updating Information About Institution Characteristics

The primary source of information about an institution's level and control and enrollment size was the 1987-88 IPEDS IC and Fall Enrollment (EF) files. These data formed the basis for the

sampling of institutions. The data were further updated for sampled institutions during an initial request for participation (postcard and CEO) and in subsequent contacts (for institution background data verification and data collection scheduling). Institutions that were misclassified, or had changed status between 1987-88 and 1988-89, were corrected. This process ensured that students were placed in the proper sampling stratum and sampled at the appropriate rate. Less than 3 percent of the institutions reported a different institutional classification than the original IPEDS values.

An associated issue was the verification of enrollment counts, which tended to be more problematic. To deal with this problem, institution-specific enrollment tolerance bounds were established, for each student level. If the actual enrollment for any student level fell outside of the tolerance bounds (+/- 20 percent of the expected enrollment), then the institutions were contacted to verify the completeness and accuracy of the lists. There were several lists that fell above or below the boundaries set; however, after verifying the enrollment counts and the specifications for constructing the lists, the problem was solved by replacing the lists with correct enrollment lists or changing the sampling rate. Once the actual enrollment was verified, the nominal sampling rate, for each student level, was reviewed. If the actual student level enrollments are such that the total actual within-institution student sample size fell below 12 (for August and October) or 10 (for February and June), one or more of the nominal student level rates were changed to obtain the minimum total actual sample size of 12 or 10.

Enrollment List Receipt and Processing

Because undergraduate, graduate and first-professional students were sampled at different rates, institutions were requested to supply separate lists for each level when possible. There were some institutions that could not supply lists in this format (two in the fall sample), or due to time constraints, could not provide a list before the scheduled visit. In both cases, procedures were developed to handle the situation.

In the first case, institutions sent one list with more than one student level with a variable to designate the student's level and/or degree type and not sorted by level. These lists were referred to as "mixed lists" and the sampling was done in two or three stages depending on the number of levels combined in the list. The first stage of sampling began with the student level with the largest within-institution sampling rate and then a sub-sample of the other student level(s) was selected. When the student level was not present on the list, the sample was drawn at the undergraduate level rate.

In the latter case, whenever an institution could not send a list or could not send a list in a feasible amount of time with respect to field scheduling of interviewer visits to institutions, the

interviewer in the field did the sampling.[3] Once the interviewer obtained the enrollment list(s), phone contact was made with a statistician at Westat to verify the actual enrollment, make changes in the sampling rate (if necessary), and draw the sample. There were 25 institutions that required in-field sampling and all were during the October and/or August sampling periods. After completing the sampling procedure, the interviewer verified whether the list(s) used for sampling would be mailed to Westat for further processing. It was of vital importance that we receive the October list from those institutions that would be participating in subsequent sampling periods. The October list was considered the master list, and without it the unduplication process could not be completed. Of those institutions that did not send lists, none participated in the February and June sampling.

Defining Nonfall Sampling Dates

The nonfall sampling dates were spaced in approximately three to four month intervals such that the vast majority of students, including those who might be enrolled in short three-month courses, were included in the survey. These sampling dates, however, were less familiar to the institutions than the fall date (October 15) which was used by other data collection programs such as IPEDS. To avoid the problem that the sampling date might fall between school terms, and thus no students enrolled on a particular sampling date, institutions were told to report the student enrollment in close proximity to the sampling date (i.e., on or about the specified date).

Institutional Partial Response for Nonfall Sampling

Institutions that participated during the fall sample period but refused to participate during one or more of the other sample periods were labeled as partial respondents. There were 75 partially responding institutions and non-response adjustment factors were computed for these institutions. The non-response adjustment factor, for time period t, is the sum of the enrollment (defined at the point of sampling) in the eligible institutions divided by the sum of the enrollment in the eligible and participating institutions. The non-response adjustment factors were constructed separately in classes defined by institution level and control.

It should be noted that some institutions had only one registration during the year or they had no enrollment for a particular sample period. For instance, some institutions did not have summer programs. These institutions were considered eligible and participating. On the other hand, institutions that had nonfall enrollment but refused to participate in the fall sample were not contacted for nonfall sampling and were labeled as non-responding institutions.

---

[3]These institutions were contacted to re-iterate the request until it was no longer feasible.

CHAPTER 4. INSTITUTIONAL RECORDS DATA COLLECTION

In this chapter, we describe the process of collecting student financial aid and other information at our sample of 1,130 institutions. Data on individual students was obtained from registrar and financial aid offices. Our field staff were sent to each school, where they recorded student data on a *record abstract*. To organize their activities, field staff used an *institution checklist*. The next two sections of this chapter describe the design of these two forms. Following that is a description of the scheduling process, field staff recruitment and training, field visits, and receipt control.

4.1 Record Abstracts

The Record Abstract was the main data collection instrument for institution records data collection. In designing the 1990 Abstract, two often competing goals had to be considered. The first goal was continuity with the 1987 NPSAS. To use the data in time-series analyses, it was essential that similar data be collected. Second, the Record Abstract needed to be revised and updated to reflect the dynamic nature of the financial assistance programs. In revising the Record Abstract, balancing these two goals was an overriding factor. A further consideration in revising the Abstract was to be sure the instrument served the objectives of the 1990 NPSAS which were to:

- Obtain student characteristics and periods of enrollments for students enrolled in postsecondary education at any time in the 1989-90 NPSAS school year (July 1, 1989 through June 30, 1990)..
- Obtain demographic characteristics and locating information for all sampled students for the 1989-90 NPSAS school year.
- Obtain enrollment information for all periods of postsecondary attendance for all sampled students for the 1989-90 NPSAS school year.
- Obtain all financial aid awards recorded in institutional records, including central and departmental offices.
- Obtain any budget data used in the determination of eligibility for financial aid, including Congressional Methodology and Pell allowable costs.
- Obtain total financial aid awards received by these students for attendance during the 1989-90 NPSAS school year.

- Obtain information used to determine each student's eligibility for financial aid for the 1989-90 NPSAS school year.

To meet these multiple objectives, considerable revision of the 1987 NPSAS Record Abstract was necessary. For example, we dropped items referring to obsolete financial aid programs and added items referring to newly created financial aid programs. Other revisions were based on the recommendations of the federal Steering Committee and the Record Abstract Working Group, a subgroup of the Steering Committee. Final revisions involved changes to both the structure and content of the Record Abstract.

One key change was to group questions by content and likely source. This reformatting served to simplify the data collection by creating a more natural flow. It also guided the data collectors to a more thorough search of all possible sources of data.

The first major grouping of items included those whose source was most likely the office of the registrar or admissions. Student- and parent-locating information (items 1 through 20), student demographics (items 21 through 26) and enrollment status (items 27 through 38) were identified under these headings and were grouped at the front end of the Abstract. The student- and parent-locating information appeared on the first and last pages of the Record Abstract. This data was recorded first and returned to Westat immediately for processing. The collection and early return from the field of student-locating information was a critical factor in completing the student and parent telephone interviews.

The next major groups of items were those having to do with financial aid information. Financial aid award information was grouped according to its source: Federal, State, Institutional, VA/DOD, Other and Graduate or First-Professional. This layout encouraged a more thorough review of all financial aid information including the most underreported awards in the 1987 NPSAS, Graduate/First-Professional and VA/DOD awards. In this section and in the following budget section, a new multiple-column format was used, because many schools have multiple award periods.

The next section on the form contains need analysis and budget data. Based on recommendations from financial aid administrators and the Federal Steering Committee, information on the Institutional budget and Expected Family Contribution (EFC) and a Congressional Methodology budget (CM), if different from the Institutional budget, was collected along with the Pell budget and Student Aid Index (SAI). The rationale behind this revision was that the Institutional budget better represented the student's true cost of attendance.

The final section of the Abstract contained data from financial aid applications. The source for this data was the financial aid office. Data collectors first identified which application forms were available in the student's file. After identifying the form(s), a skip pattern was triggered allowing the data collector to abstract information for just that form, ignoring the others. For the most common aid application forms, the Abstract contained a template containing the data elements in the same order in which they appear on the application documents.

4.2 Institution Checklist

The Institution Checklist was a multi-purpose form used by the data collector to compile school-level data. The Checklist served as a guide for the initial meeting with the institution coordinator and as a guide throughout the conduct of data collection at the school. The 17-page Checklist contained the following information:

- Preprinted information about the school,
- Sample sizes by student level,
- A reference for the specific location and nature of various record keeping systems at the school,
- School policies and procedures related to the calendar system, grading system, attendance status, tuition, and financial aid information,
- Documentation of any circumstances or procedures outside the realm of standard field procedures as discussed in training, and
- Name of the data collector and the dates of collection.

The Checklist for NPSAS 1990 was based on its 1987 counterpart. It was, however, modified in length, degree of complexity, and content. The length of the Checklist was increased by four pages over the 1987 instrument. All questions about school policy and procedures were asked about each level of student rather than as general policy questions. This change resulted in information that was more accurate and more useful to the data collector. The detail in each item was reduced. This reduced the burden on the school coordinator while slightly increasing the work of the field staff.

The completed Institution Checklist became a road map to be followed during data collection at the institution. It directed the data collector to various sources and provided information essential to the completion of the Abstract. Also, the Checklist was used by processing staff, upon receipt of completed Record Abstracts, to seek clarification of data problems, such as missing data and unclassified awards.

Detailed Look at the Checklist

The first page of the Checklist was created from the receipt control system and provided contact and sampling information about the institution. All information provided on this sheet proved to be useful to the data collectors. Of particular value was the variable named "SPECIAL PROCEDURES" which briefly summarized any special arrangements made with the school during recruiting and scheduling. Obviously, advance knowledge of the arrangements made for better-prepared data collectors.

The next section of the Institution Checklist (pages 2 - 9) provided space to record the source and location, by student level, for those records to be used to complete the Record Abstract. This was necessary because at most traditional institutions the information requested in the Record Abstract was likely to be found in different offices. For example, financial aid for graduate students is usually administered and stored in a separate graduate financial aid office or in individual financial aid offices within graduate departments. This is in contrast to financial aid for undergraduates which is likely to be processed by a central financial aid office.

The next section of the Checklist, "Institution Information" provided questions on school policy and procedures to be asked of the school coordinator by the data collector. The questions were keyed to specific items in the Abstract, and were used as a reference when completing certain sections of the Record Abstract. As necessary, the questions were asked of each level of student--undergraduate, graduate, and first-professional.

The next part of the Checklist focused on the school's aid programs. Data collectors discussed with the school coordinator the various assistance and award programs in which the school participated. This discussion alerted the data collector to the type of awards likely to be encountered in the student records. It also provided an opportunity for clarification for awards not easily categorized. This format was developed to reduce the use of the "other, specify" response by data collectors in the Record Abstract.

Data collectors used the last section of the checklist to record procedural decisions made in the field and other documentation of their data collection visits.

4.3 Scheduling Institution Visits

Institutional coordinators were mailed a scheduling package which included a cover letter to coordinators and a Data Collection Scheduling Information sheet (Exhibit 4-1). This scheduling sheet reminded the coordinators of enrollment list requests, identified the institution's enrollment counts and its approximate fall sample sizes, and listed a proposed visit date for data collection. These dates were based on an ideal schedule constructed in advance of this mailing. Also, included in this package were final draft copies of the Institution Checklist and the Record Abstract.

Exhibit 4-1

1990 NATIONAL POSTSECONDARY STUDENT AID STUDY (1990 NPSAS)
DATA COLLECTION SCHEDULING INFORMATION

INSTITUTION LOCATING DATA:

Listed below are the coordinator name and institution locating information we have listed for this institution.

COORDINATOR NAME: _____ ID: _____

INSTITUTION NAME:_____

STREET:_____

CITY:_____

STATE:_____ ZIP:_____

IF ANY OF THIS INFORMATION IS INCORRECT PLEASE CALL THE NPSAS INFORMATION LINE (1-800-937-8288.)

ENROLLMENT LIST REMINDER:

In September you received a packet of materials in a blue folder labelled "STUDENT ENROLLMENT LIST REQUEST". This packet contained student eligibility information, examples of enrollment list formats, and dates by which we would like to receive the lists. In order for us to maintain the current study schedule, it is very important for us to receive your enrollment lists on or before the dates specified.

IF YOU ANTICIPATE ANY DELAYS IN PROVIDING ANY OF THE REQUESTED LISTS PLEASE CALL THE NPSAS INFORMATION LINE (1-800-937-8288) AS SOON AS POSSIBLE.

THE STUDENT SAMPLE:

The previous packet also provided our estimates of your institution's 1989 Fall enrollment and asked that you inform us if those figures differed significantly from your estimates. Shown below are the estimates of your institution's 1989 Fall enrollment that we currently have on our file as well as the approximate number of students we will be selecting.

1989 FALL ENROLLMENT ESTIMATES	APPROXIMATE NUMBER OF STUDENTS TO BE SELECTED
UNDERGRADUATE:_____	_____
GRADUATE: _____	_____
FIRST PROFESSIONAL:_____	_____

FALL DATA COLLECTION SCHEDULE:

In early December, Westat and the National Association of Student Financial Aid Administrators (NASFAA) will train approximately 150 field representatives. One of these representatives will then visit you to review the Institution Checklist and complete a Record Abstract for each sampled student. Shown below is the date on which we would like to have one of our field representatives begin data collection at your institution and also an estimate of the number of days it might take him/her to complete the abstracts.

VISIT START DATE: _____ EXPECTED LENGTH OF VISIT: _____

IF IT WILL NOT BE POSSIBLE FOR US TO VISIT YOUR INSTITUTION ON THE DATE SPECIFIED ABOVE, PLEASE CALL THE NPSAS INFORMATION LINE (1-800-937-8288) AND WE WILL TRY TO RESCHEDULE THE VISIT AT YOUR CONVENIENCE.

All participating institutions were then called to confirm or revise the suggested date. These follow-up phone calls began on October 30, 1989 and continued for approximately two weeks. After the follow-up call was made, scheduled visit dates were confirmed with a letter to the school. The established schedules were flexible, though, and some were revised throughout the field period.

4.3.1 Creating an Ideal Master Schedule

All participating institutions and all institutions with a status of "initial refusal" were assigned to a specific geographic cluster referred to as a field group. The field group represented the work load of one data collector across the entire field period (December 1989 to March 1990). The intent of assigning an institution to a particular field group was to reduce the amount of travel within a particular assignment. For the 1990 NPSAS, 156 field groups were created. Each field group contained between one and twelve institutions of various sizes, level and control.

Using scheduling calendars (Exhibit 4-2), all institutions within a particular field group were assigned a preferred visit date. The dates assigned to each institution were based on travel distance between schools, the most feasible travel plan for the entire field group, and the expected work load at each institution. Work load was projected based on the 1987 NPSAS and 1990 Field Test average of 25 completed Record Abstracts per data collector per day. The scheduling calendars identified each institution by name and Westat ID number. They also indicated the starting and anticipated ending date for data collection at the institution. These calendars were used as the basis for scheduling the institution visits. The calendars were revised and amended throughout the field period.

4.3.2 Scheduling Calls

Receipt of the scheduling package by the institution coordinator resulted in phone calls to Westat operations staff from about 30 percent of the participating schools. All of these calls were to discuss the suggested visit date. Those schools not responding to the mailing were telephoned to confirm the suggested visit dates.

Exhibit 4-2

Scheduling Calendar

February 1990

SUN	MON	TUE	WED	THU	FRI	SAT
				1 1063421 Ind Bty Acad. ---------->	2	3
4	5 1062456 Old South Univ. ---------->	6	7	8	9	10
11	12 1062456 Old South Univ. (cont.) ---------->	13	14	15	16	17
18	19 1063284 Winston Jr. Col. ---------->	20	21	22	23 1063974 Ross Vo-Tec Inst. ----->	24
25	26 1063974 Ross Vo-Tec Inst. (cont.) ----->	27	28			

January 1990

S	M	T	W	T	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

March 1990

S	M	T	W	T	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

The scheduling calls began on October 30th and 98 percent were concluded by November 30th. The remaining calls were completed while the field period was under way.

Because of requests by institutions to reschedule, it became necessary to create approximately ten additional field groups. When an institution's scheduling request could not be met by its field group, it was assigned to one of the new field groups. In this way all schedule requests were fulfilled.

4.3.3 Scheduling Database

The data obtained about procedural aspects of the field visit and confirmed visit dates were recorded on the Institution Scheduling Screening Guide. The information from the guide was entered into our main institution receipt control system. Reports on this scheduling phase of the study were generated from receipt control on a weekly basis. The information obtained during the scheduling contacts was used to produce two forms for the data collector:

- a Field Assignment Summary Sheet (Exhibit 4-3) which provided scheduling dates and times for all scheduled schools in a particular field group, and
- an Institution Information Sheet which provided information specific to the data collection to be completed at each school.

4.3.4 Confirmation Packets

As scheduled visit dates were confirmed, a confirmation package was assembled and mailed to each coordinator. The mailing began the first week of December 1989 and continued through the second week of January 1990. This package included a letter specifying the date(s) of the visit and its expected length and any special procedures that had been agreed upon. The letter provided the phone number for the NPSAS Information Line. It also identified a person at Westat to be contacted if there were any problems with or changes to the agreed upon dates. Included in the confirmation package was the final version of the Record Abstract and the institution copy of the Student Sample Listing Sheet which identified the students selected for participation in the study.

Exhibit 4-3

NATIONAL POSTSECONDARY STUDENT AID STUDY
FIELD ASSIGNMENT SUMMARY SHEET
FIELD GROUP _____

WESID	INSTITUTION NAME	CITY	INSTDATE	VISIT LENGTH	SAMPSIZE UGRAD	GRAD	PR

[TO BE PRODUCED BY RECEIPT CONTROL SYSTEM FOR EACH FIELD GROUP]

4.3.5 Re-scheduling

Approximately 25 percent of the participating institutions required a schedule change during the fall data collection. About seven percent of participating schools required two or more schedule changes. Schedule changes were necessary for a variety of reasons and usually were made at the school's request.

As scheduling changes took place, multiple updates were made. In the institution file, all hard-copy forms having a scheduling date on them were updated. The scheduling variables in the institution receipt control file were also updated to reflect the change. Finally, a new confirmation letter was generated and sent to the institution coordinator. Copies of this letter were also sent to the assigned data collector and the field supervisor and a copy was kept in the institution file. Field supervisors were also alerted of schedule changes via phone calls from operations staff. They quickly relayed this information to the appropriate data collectors.

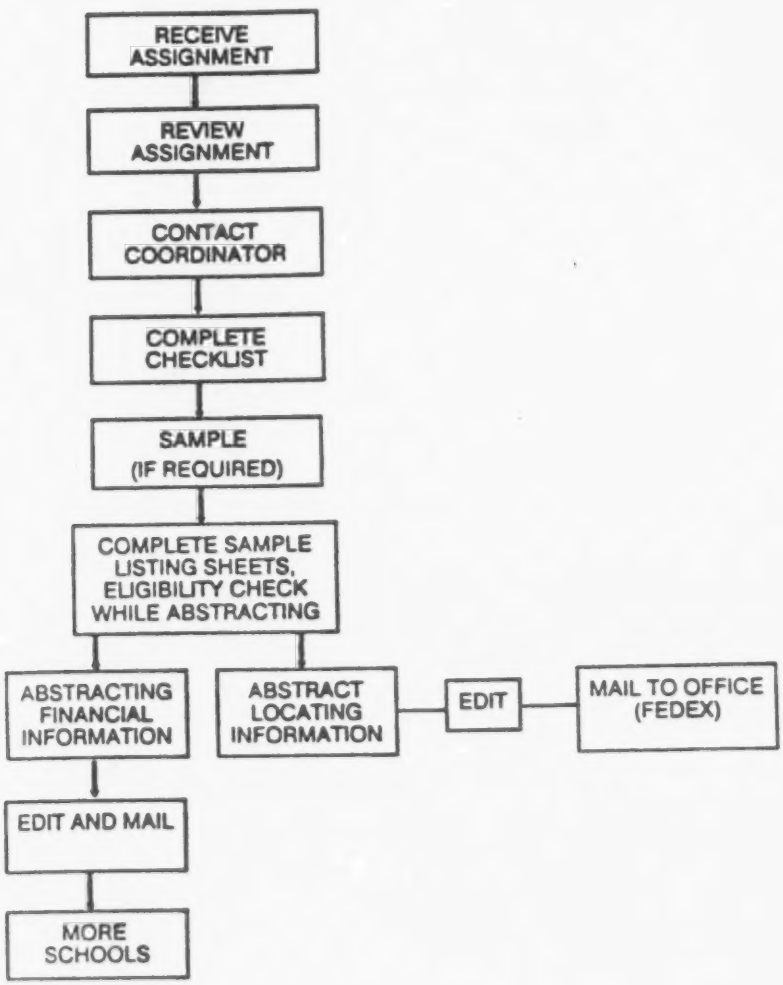
Over 75 percent of the data collection visits took place during January and February 1990. Over half of the visits were scheduled between January 22nd and February 23rd.

4.4 Field Data Collection

The field data collection for the 1990 NPSAS sought school-reported student-level data as well as locating information (local and permanent addresses and phone numbers) for sampled students and their parents. The field periods for data collection covered two periods of time. The first field period, to collect fall records data for students sampled from summer and fall 1989 enrollment lists, began December 7, 1989 and over 99 percent was completed by March 17, 1991. The second field period had a dual purpose: 1) updating the fall records data previously collected, and 2) collecting nonfall records data for students sampled from winter/spring and summer 1990 enrollment lists. This data collection was conducted beginning August 19, 1990 and continued through November. Exhibit 4-4 provides the flow of data collection activities for the study.

Exhibit 4-4

DATA COLLECTORS' TASK



Other Data Collection Materials

The data collection instrument for both fall and nonfall records was the Record Abstract (see section 4.2.1). Additional key materials used during records data collection were the Student Sample Listing Sheet(s) and the Institution Checklist (see sections 3.4 and 4.2.2, respectively). Several additional materials, including disclosure notices, institution labels, and training information were developed.

Disclosure Notice

For many schools, confidentiality of student records was a particularly important issue. For this reason, the disclosure notice (Exhibit 4-5) was developed. The notice, briefly describing the study and citing the regulations which permit access to student files without student notification or permission, was inserted in each accessed student file at schools that had requested that this be done. Approximately 66 percent of participating institutions requested use of the disclosure notice.

Institution Labels

For each participating institution, a set of preprinted labels was generated. The labels consisted of the school name and a unique seven-digit identification number (ID). The first three digits represented the primary sampling unit (PSU) in which the institution was located, the next three digits were a unique school identifier, and the last digit was a check digit. Labels were also produced for each sampled student at the institution and were to be placed on the completed Record Abstracts. These labels included the sampled student's name and a thirteen-digit student ID. The first seven digits represented the student's school and were in fact the same as the institution ID. The next digit represented the sample month (August, October, February or June) from which the student was sampled. Digit 9 represented the student's level, undergraduate, graduate or first-professional. The next three digits represented the unique student ID and the last digit was a check digit. Both types of labels contained bar codes so incoming data could be quickly and accurately receipted using an optical scanner.

Exhibit 4-5

DISCLOSURE NOTICE
1990 NATIONAL POSTSECONDARY STUDENT AID STUDY

Information from student records for this individual(s) has been provided to Westat, Inc., agent for the U.S. Department of Education. This disclosure statement fulfills the requirements of 34 CFR 99.32(a)(6) pursuant to the Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. 1232g).

This individual is included in the sample for the 1990 National Postsecondary Student Aid Study and data from these records will be combined with others into statistical summaries and tables. No individually identifiable information will be released in any form.

Field Group Assignment Sheet

This computer generated sheet listed the institutions assigned to a particular field group and information about the schedule and work load at each school. The specific variables on this form included:

- the field group number, the school ID,
- the name of the institution,
- the city in which it was located,
- the date on which data collection was scheduled to begin, and
- the sample size for each level of student enrolled at the institution. This sheet summarized the data collector's assignment over the course of the field period.

NPSAS Field Schedule Calendar

An updated Field Schedule Calendar (see section 4.3.1) was produced for each field group. The calendar listed by name and WESID all schools in a particular field group having confirmed visit dates. The calendar listed the beginning date of data collection for each scheduled school and indicated the amount of time scheduled for the completion of data collection. The calendar format provided good visual reference for the expected work load.

4.5 Supervisor Recruitment and Training

The requirements of the records data collection required an efficient, well-trained and organized supervisory staff to facilitate field operations. Developing such a staff required organizing field groups into geographical regions, hiring and training field supervisors and developing a field reporting system.

Field Group Assignment

As discussed earlier, a field group represented a logical grouping of institutions based on geographical location and expected length of data collection. Each field group represented one data collector's assignment. Each field group was assigned a three digit ID; then, based on their location, they were assigned to one of eight regions. Each region, except for Puerto Rico, represented between 18 and 25 field groups encompassing 3 to 13 states. The size of each region varied considerably because of population densities and subsequent institution and student sampling. Puerto Rico, because of its location, was designated its own region. The supervisor in Puerto Rico was responsible for eight data collectors.

Recruitment

Staffing needs were based on our previous experiences with large-scale field operations. These indicated an optimal ratio of approximately twenty field data collectors per supervisor. Based on the expected number of participating schools and the work load associated with these schools, we estimated that the study would require approximately 170 data collectors. This necessitated the hiring of eight supervisors.

Seven of the supervisors were distributed across the country (California, Florida, Maryland, New Jersey, Texas and Wisconsin) and one supervisor was hired for assignments in Puerto Rico. This structure made communication between field staff and supervisors more efficient (time zones became a non-issue) and more cost-effective (the cost of telephone communication between supervisors and field staff located in the same areas of the country was less). Also, as supervisors were likely to know potential data collectors in their area, recruiting efforts also benefitted from this structure.

All the supervisors hired for the 1990 NPSAS were experienced as supervisors on other education studies, such as, the High School Transcript Study, the National Assessment of Educational Progress and the 1990 NPSAS Pilot Test.

Supervisor Training

Supervisors attended a three-day training session (November 27, 28 and 29) at the Westat home office. A supervisor's field manual was prepared and used as the basis for this training. This training provided background information about the study, technical and procedural information about the data collection (with a strong emphasis on the Institution Checklist and the Record Abstract) and administrative information, including training on the Automated Survey Control System (ASCS) for field reporting. The training involved interactive lecture as well as practice sessions with selected training exercises. The supervisors were also trained for their expected roles as assistants to the lead trainer during data collector training.

The Automated Survey Control System

Westat programming staff developed an Automated Survey Control System (ASCS) to track the progress of field work for the institutional records data collection. The system operated on compact IBM-compatible personal computers located in each field supervisor's office. Each PC was connected by phone lines to the Westat main office which would allow direct transmission of reports from the field.

ASCS was a menu-driven system that created and used two databases. One database kept records on data collectors' expenses and production. The second data base recorded information on institutions including name, WESID, type, control, scheduled visit date and date completed. ASCS reports provided a general picture of the rate of completion in the field. It included production reports by field group/data collector, numbers of Record Abstracts completed, and the date on which completed data was mailed from the field to Westat.

4.6 Data Collector Recruitment and Training

Field labor requirements were estimated to be approximately 170 field data collectors for a 12-week field period. Training was held during December and January at four different sites, including Puerto Rico, to allow adequate training of all staff and to reduce travel costs to and from training.

Recruitment

As soon as they were hired, field supervisors began recruiting data collectors. On October 16, 1989 the field director mailed to supervisors a package of materials which included:

- a memo providing a brief project description and recruiting information;
- a list of the states assigned to each of the eight NPSAS field regions;
- a list from Westat field files of the available data collectors in each region;
- copies of incoming data collector phone calls inquiring about project work;
- copies of the field scheduling calendars indicating distribution and level of work in each region; and
- a list of all sampled schools in each region.

Most of the data collectors hired for the 1990 NPSAS came from Westat interviewer files. About 70 percent of those hired had previous Westat experience as interviewers or data collectors and approximately half of these had worked on the 1987 NPSAS.

Recruiting of data collectors for the 1990 NPSAS went smoothly, in large part due to the experience of the field supervisors. By October 31, 1989 supervisors had hired 50 percent of the required staff and by Thanksgiving, 98 percent of the staff had been hired.

Most data collectors were assigned a specific field group and thus had some idea of their work load over the entire field period. In seven of the eight regions, one or two data collectors were hired as travelers/back-ups. They received no specific set of assignments but were hired with the understanding that they would be available to work at any institution in the region where assistance was required. Travelers were not hired in Puerto Rico because the work load was considerably less there than in other regions. Staff hired in Puerto Rico were willing to travel to lend assistance as necessary.

The Field Manual

An Institution Data Collection Field Manual served both as the basis for data collector training and as a reference guide for the data collectors during the field period. The manual was based on both the 1990 NPSAS Pilot Study Field Manual and the 1990 NPSAS Supervisor Manual. The manual was developed over a two month period (mid-September to mid-November) by Westat project staff and NCES. The manual included background information, a review of advance activities (from institution enlistment through scheduling field visits), discussion of confidentiality and professional ethics, the issue of quality control, and administrative procedures. The manual included exhibits of all advance materials as well as of all field forms.

The two most significant sections of the field manual were those providing descriptions of and procedures for using the Institution Checklist and the Record Abstract. In each of these sections, copies of the instruments were included along with item-by item specifications on completing each item. Notes were included in the specifications to highlight certain suggestions or warnings based on previous NPSAS experience. The section covering the Record Abstract included two examples of data collection.

Other Training Materials

Numerous other materials were developed for use in training by the field staff. They included:

- a scripted training guide;
- training exercise; and
- role-playing activities.

Training

A training walk-through was held at Westat on November 15, 16 and 17. The walk-through included all training staff and involved a complete presentation of data collector training and a discussion of the trainers' role and responsibility at training. The walk-through also included practice sessions with selected training exercises.

Four training sessions, each lasting four and one half days, were held for data collectors before starting the field work. In November and December, three sessions were conducted as follows:

November 30 - December 4	Washington, D.C. Regions 1, 2 and 3
December 7-11	New Orleans, LA Region 4
December 14-18	Los Angeles, CA Regions 5, 6 and 7

The fourth session, for region 8, Puerto Rico, was held in San Juan, Puerto Rico from January 15 through January 19, 1990.

The sessions began with an introductory welcoming meeting including opening remarks by the Westat corporate officer, the national field director and, at the Washington and Los Angeles sessions, key NCES staff. The remainder of training was patterned after the organization of the Field Manual (Exhibit 4-6). Each day of training focused on one particular component of the study. Day one provided background information about the study and the contacts with participating institutions to the field visits. The second day focused on the field forms to be used for data collection with special emphasis on the Institution Checklist. The third and fourth days of training provided an overview of financial aid and the completion of the Record Abstract. The final day of training was used as a final review and to discuss administrative tasks related to the study.

The training sessions were led by Westat project staff who were supported by the field supervisors. NASFAA staff members or a financial aid officer were available as resources during the two days of training in which the Record Abstract was covered.

Exhibit 4-6

TABLE OF CONTENTS

Chapter			Page
1	INTRODUCTION AND OVERVIEW		1-1
	1.1	Background Information of the 1990 National Postsecondary Student Aid Study	1-1
	1.2	Purposes and Goals	1-2
	1.3	The Design	1-3
	1.4	The Institution Sample	1-4
	1.5	The Student Sample	1-5
	1.6	The Parent Sample	1-6
	1.7	The Study Schedule	1-6
	1.8	Overview of Data Collectors' Tasks	1-6
2	REPRESENTING WESTAT AND THE DEPARTMENT OF EDUCATION		2-1
	2.1	Explaining and Representing the Study	2-1
	2.2	Attitude, Conduct and Professional Ethics	2-1
	2.3	Confidentiality	2-2
3.	PREVIOUS CONTACTS WITH INSTITUTIONS		3-1
	3.1	Advance Packets	3-1
		3.1.1 Telephone Followup	3-1
		3.1.2 Refusal Conversion	3-11
	3.2	Coordinator Packets	3-11
	3.3	Scheduling Packets	3-15
	3.4	Field Visit Scheduling	3-15
	3.5	Confirmation Packets	3-17
4	PREPARING FOR INSTITUTION VISITS		4-1
	4.1	Overview of Field Materials	4-1
	4.2	Planning the Field Visits	4-3
	4.3	Confirming Appointments	4-3
	4.4	Documenting Changes/Problems	4-6
5	CONDUCTING FIELD VISITS		5-1
	5.1	Using the Forms	5-1
		5.1.1 The Field Group Assignment Sheet	5-1
		5.1.2 NPSAS Field Schedule Calendar	5-3
		5.1.3 Institution Information Sheet	5-3
		5.1.4 Institution Checklist	5-6

Exhibit 4-6 (continued)

TABLE OF CONTENTS (continued)

Chapter			Page
	5.1.5	Student Sample Listing Sheet	5-7
	5.1.6	The Record Abstract	5-10
	5.2	Abstracting Record Data	5-11
	5.2.1	Locating Best Data Sources	5-12
	5.2.2	Handling Faded and Illegible Documents	5-13
	5.2.3	Handling Incomplete Documentation	5-13
6		THE INSTITUTION CHECKLIST	6-1
	6.1	Overview	6-1
	6.2	Item-by-Item Specifications for the Institution Checklist	6-5
	6.2.1	The Institution Information Sheet (Page 1)	6-5
	6.2.2	Records Source Information (Pages 2-9)	6-13
	6.2.3	Institution Information (Pages 10-13)	6-25
	6.2.4	Record Abstract Categories (Page 14)	6-33
	6.2.5	Documentation (Pages 15 and 16)	6-37
7		THE RECORD ABSTRACT	7-1
	7.1	Introduction to the Record Abstract	7-1
	7.2	Overview of the Record Abstract	7-3
	7.3	Background Information for the Record Abstract	7-4
	7.4	Sources of Record Abstract Information	7-11
	7.5	Two Examples of Record Abstract Data Collection	7-20
	7.6	Question-by-Question Specifications	7-23
8		QUALITY CONTROL	8-1
	8.1	Importance of Field Edit	8-1
	8.2	NPSAS Field Edit	8-3
	8.3	Supervisor Review and Feedback Channels	8-3
	8.4	Contact Persons for Problem Resolution	8-6
9		ADMINISTRATIVE PROCEDURES	9-1
	9.1	The Time and Expense Report (T & E)	9-1
	9.1.1	General Consideration for Completing the T & E	9-3
	9.1.2	Specifications for Completing the T & E	9-4

Exhibit 4-6 (continued)

TABLE OF CONTENTS (continued)

Chapter			Page
9.2	Travel Outside Your Data Collection Area		9-7
	9.2.1	Approval of Out-of-town Arrangements	9-7
	9.2.2	Air Transportation	9-7
	9.2.3	Lodging	9-8
	9.2.4	Personal Automobile	9-10
	9.2.5	Local Public Transportation	9-10
	9.2.6	Car Rentals	9-11
	9.2.7	Telephone Calls	9-11
	9.2.8	Miscellaneous Expenses	9-12
	9.2.9	Supporting Receipts	9-12
	9.2.10	Reporting Expenses on Westat Trip Expense Form	9-12
	9.2.11	Travel Advances	9-13
9.3	Reporting to the Home Office		9-13
9.4	Mailing Materials		9-14

LIST OF EXHIBITS

Exhibit		
2-1	Assurance of Confidentiality Pledge	2-4
3-1	NCES letter (advance packet)	3-2
3-2	NPSAS Study Summary Pamphlet	3-3
3-3	Information Sheet	3-7
3-4	Authorization and Confidentiality Statement	3-8
3-5	Endorsement	3-9
3-6	1990 NPSAS Postcard	3-10
3-7	Institution Coordinator Letter	3-12
3-8	Institution Background Data Verification Sheet	3-13
3-9	1990 NPSAS Student Enrollment List Request	3-14
3-10	1990 NPSAS Coordinator Scheduling Letter	3-16

BEST COPY AVAILABLE

Exhibit 4-6 (continued)

TABLE OF CONTENTS (continued)

LIST OF EXHIBITS (continued)

Exhibit		Page
3-11	NPSAS Scheduling Confirmation Letter	3-18
4-1	1990 NPSAS Field Forms	4-2
4-2	Field Assignment Summary Sheet	4-4
4-3	Scheduling Calendar	4-5
4-4	Confirmation Items Checklist	4-7
5-1	Field Visit Checklist	5-2
5-2	Student Sample Listing Sheet	5-8
8-1	Locating Information Edit Checklist	8-4
8-2	Record Abstract Edit Checklist	8-5
8-3	Interviewer Edit Feedback Form	8-7
9-1	Time and Expense Report	9-2
9-2	Trip Expense Report	9-9
9-3	NPSAS Shipping Transmittal Form	9-15

LIST OF APPENDICES

Appendix		
A	1990 NPSAS Field Forms	A-1
B	1990 NPSAS Contact List	B-1
C	Output Documents	C-1
D	Record Abstract Reference Guide	D-1
E	State and Province Abbreviations	E-1
F	Glossary of Terms	F-1

The two basic training techniques used to cover field procedures and use of the field forms were interactive lectures to the groups presented by the Westat trainers and various exercises and role-play activities.

4.7 Field Visits

The field period for fall records data collection involved over 1,000 visits to institutions, ninety percent of which were completed from December 1989 through February 1990. The data were collected by one to three data collectors over field visits lasting from a day to five weeks, depending on the work load at an institution.

Field Assignments

The last part of the last day of training was used to distribute and review field assignments with the data collectors. The assignment packages included several field group-specific and institution-specific forms. Each package included a scheduling calendar which indicated the starting date and the duration of the visit for all scheduled schools in an assignment.

Data collectors reviewed the assignment packages with their supervisors at the end of training. This provided a check against materials expected and materials received as well as a early warning of any scheduling conflicts.

The forms included in the field assignment packages were updated, as necessary, throughout the field period. For example, when a confirmed visit date was changed, field staff were alerted, by phone, immediately. Also, as enrollment lists were received and the student sample selected, data collectors were sent the multiple copies of the sample listing sheets.

Confirming Appointments and Preparing Materials

Data collectors telephoned institution coordinators to introduce themselves and to confirm the visit date and arrangements. Field staff were instructed to make these calls at least three days before the visit. This provided the opportunity for the data collector to review the arrangements necessary to ensure that the field visit ran smoothly.

Before going to an institution, the data collector organized the institution-specific materials for the visit and counted out the necessary number of Record Abstracts and other supplies.

Field Visits

Field visits lasted from one day to five weeks. While the majority of institutions were willing to provide the amount of time requested, some were not. For this small number of cases, a faster data collection was accomplished by sending multiple data collectors to those schools.

Field Edits

Data collectors were asked to edit completed materials, either before leaving an institution, or before shipping materials to Westat. Two forms were developed to assist the field editing process, the Locating Information Edit Checklist and the Record Abstract Edit Checklist. The checklists provided a spot check of the work done by the data collector. This technique served to identify any large-scale problems.

Supervisor Editing and Reporting

Supervisors scanned the work for overall completeness and the Record Abstracts for missing data, adherence to skip patterns, and legibility. Any major errors were brought immediately to the data collector's attention. Otherwise, this cursory edit of the data collector's work was relayed to him during his telephone report to the supervisor.

The supervisors reported on completion rates and expenses in their region on a weekly basis via the Automated Survey Control System. Telephone reports from the field staff to the supervisors provided information to the supervisors which was then loaded into ASCS.

Schools That Did Their Own Abstracting

For the fall records data collection, only two schools asked to do the data collection themselves. One wanted to do only the financial aid section of the abstract. Each school was sent:

- multiple copies of forms (Sample Listing Sheets, Record Abstracts and student labels),
- a chapter from the field manual, and
- a cover letter, including the NPSAS technical assistance number.

Technical Support

Technical support for the field staff was provided by various members of the Westat project staff. Field staff were instructed to use the 800 NPSAS Information Line to call the home office with technical questions only after first attempting resolution through the appropriate field supervisor. Documentation of these calls was recorded on the 1990 NPSAS Data Collection Technical Assistance Log. This NCR form was used to record the source of the call (data collection staff, supervisor, institution or other), the question or problem presented, and the resolution. Information from the form was used to update the receipt control system, as warranted, (schedule changes, for example). Copies of the form were then placed in the hard-copy institution file and placed in two different sets of binders, one organized by WESID and the other by date.

4.8 Receipt and Processing

Receipt Control

The 1990 NPSAS receipt control system consisted of three files: a school-level, a student-level and a batch file. Each file could be accessed for updates through a menu screen. The school-level file was used to record the receipt date of the covers of the Record Abstracts for a completed school and the entire package of forms and materials associated with a finished school. The cover pages of the Abstract were always completed first by the field staff and were sent immediately to Westat for processing so that the information which they contained could be loaded into the Student and Parent CATI systems. When school-level files were received and had been logged into the receipt system as such, the next step was to update the student-level file with both a receipt date and a batch date for both the covers and the Record Abstracts for each sampled student in preparation for sending these forms to data entry. The batch file indicated, for each batch number, the total number of forms included in the batch, all the schools included in the batch, and the date the batch was sent to and returned from data entry. Additionally, for quality control, the first and last student ID numbers in the batch were recorded, thereby making it possible, to do a quick check for all forms.

Receipt and Scan Editing

In December 1989, clerical staff were hired and trained to receive and process forms and materials arriving from the field. Training focused on using the receipt control system, the process of receipting incoming mail, and edit checks to be performed during the initial processing.

In January 1990, additional staff were hired to perform scan edits of the record abstract data entry. The tasks to be performed by this staff were explained in a training session conducted by the

project's data preparation supervisor. First, staff were to count the Record Abstracts and check them against the batch checklist sheet for completeness. Second, the Abstracts were checked to see that proper skip patterns had been followed and that the flow of the form was consistent with those patterns. Third, individual question items were examined and coded for easier processing. Any problems encountered, such as incorrect skips or items requiring calculation, were noted on problem sheets and referred to supervisors for resolution.

Coding was verified for all cases during the first week of the operation. Fifty percent during the second week and 10 percent during the remaining weeks.

Data Entry

Data entry was performed both by external contractors and Westat's data entry facilities. Double-key verification was done for all data. For each batch sent to data entry, a machine readable file was requested to allow for a higher level of control and easier verification.

When batches were returned from data entry, the date returned was entered into receipt control. Staff manually checked each returned batch being sure that all forms indicated on the batch control sheet had been returned. The batch folders were checked against the accompanying data tape labels to verify that each returned batch was represented by a tape. The tapes were then logged on a control sheet which showed the data received, the tape identification number and the numbers of the batches on the tape label.

4.9 Nonfall Records Data Collection

Records data collection at those institutions from which a student sample was drawn from February or June 1990 enrollment lists required an additional field visit. The second field period was scheduled for the last week of August 1990 through November of 1990. It involved field visits to over 500 institutions requiring a field staff of 57 data collectors and three field supervisors. The second field period involved two major data collection tasks: completion of the Record Abstract for nonfall sampled students and the update of financial aid information collected from institutions during the first field period. Records data collection for the nonfall sample was done the same way it was done in the first field period, as described above. The Records Update Task is described in detail in the following chapter.

CHAPTER 5. INSTITUTIONAL RECORDS UPDATE

This chapter describes the methods used to update the institutional records data for students whose records were sampled in October.

5.1 Records Update Form

The Student Financial Records Update Form was designed to capture the dynamic nature of financial aid award amounts over the course of an academic year. Using this form, field data collectors could verify, correct, and update the financial aid award amounts collected earlier. The records update task recorded any changes that had occurred in financial aid award amounts during the award year (see Exhibit 5-1).

5.2 The Control List

To facilitate the organization of files required to complete the updates, all institutions were provided a copy of the institution-specific control list. The control list identified by name, NPSAS student ID and, if available, school ID all eligible students sampled in August and October of 1989. The list could also be used as a check against completed update forms to be sure that all requested updates had been completed.

5.3 The Other Specify Awards Form

To verify the classification of unspecified awards and to collect further information on them, an Other Specify Award Sheet was created for each school. The form listed each award categorized as "other" on any of the abstracts completed at a school. It listed the award name, as recorded on the abstract, and requested that the school identify the source, kind (need based or non-need based) and type (grant or loan) of the award.

1990 NATIONAL POSTSECONDARY STUDENT AID STUDY
STUDENT FINANCIAL RECORDS UPDATE FORM

OMB No.
Expiration Date:

INSTITUTION: Any College	NPSAS INSTITUTION ID: 14928730
STUDENT NAME: Student, John Q.	NPSAS STUDENT ID: 149287305103
STUDENT SSN/ID: xxx-xx-xxxx	
DEPENDENCY STATUS: Independent	

1 Was this student enrolled in this institution during:

	YES	NO
Summer '90 (on or about June 15, 1990)	1	2

2. FINANCIAL AID FROM RECORD ABSTRACT

Please review the information below

Please use this space to confirm, clarify or correct the information shown on the left.

Category	Award Name	1989-90 Award Amount	Award Period	No change	1989-90 Final Amount	Award Period
A. Federal Aid (See Instructions)	1. CWSP					
	a. Amount awarded	$500.00	9/89 to 5/90	☐	$ .00	___/___ to ___/___
	b. Amount earned				$ .00	___/___ to ___/___
	2. Pell Grant	$1,000.00	9/89 to 12/89	☐	$ .00	___/___ to ___/___
	3. Pell Grant	$1,000.00	1/90 to 5/90	☐	$ .00	___/___ to ___/___
B. State Aid (See Instructions)	1. SSIG	$500.00	9/89 to 5/90	☐	$ .00	___/___ to ___/___
C. Institutional Aid (See Instructions)	No Institutional Aid recorded for this student. To report Institutional Aid, use Section 3 on reverse side.					
D. Other Aid (See Instructions)	1. Ryan Association	$500.00	9/89 to 5/90	☐	$ .00	___/___ to ___/___

89

101

SEE REVERSE SIDE FOR ADDITIONAL INFORMATION

12

Figure 5.1

3. **ADDITIONAL AID**

Did this student receive any financial aid for the 1989-90 school year that was not listed on the previous page?

YES ..... 1 PLEASE RECORD INFORMATION CONCERNING THESE AWARDS IN THE BOX BELOW, THEN CONTINUE WITH ITEM 4

NO ..... 2 PLEASE SKIP TO ITEM 4

Award Name	Source CIRCLE ONE CODE BELOW: Federal	State	Institutional	Other	1989-90 Final Amount	Award Period
1. _____	1	2	3	4	$_____.00	__/__ to __/__
2. _____	1	2	3	4	$_____.00	__/__ to __/__
3. _____	1	2	3	4	$_____.00	__/__ to __/__
4. _____	1	2	3	4	$_____.00	__/__ to __/__
5. _____	1	2	3	4	$_____.00	__/__ to __/__

4. **COST OF ATTENDANCE**

Please review the budget information below and either confirm or update in the space provided.

A. PELL Budget

	Initial Budget Period 09/89 to 05/90	Final Budget Period ___/___ to ___/___		
	1989-90 Initial	No Change	1989-90 Final	
1. Student Aid Index (SAI)	00985	☐	_____	

B. INSTITUTIONAL Budget

	Initial Budget Period 09/89 to 05/90	Final Budget Period ___/___ to ___/___		
	1989-90 Initial	No Change	1989-90 Final	
1. Expected Family Contribution (EFC or FC) Total	_____	☐	$_____.00	
OR				
1a. Parental Contribution	$2100.00	☐	$_____.00	
1b. Student's Contribution from Income	$500.00	☐	$_____.00	
1c. Student's Contribution from Assets	0	☐	$_____.00	

69

Figure 5.1 (con't.)

BEST COPY AVAILABLE

3. **ADDITIONAL AID**

Did this student receive any financial aid for the 1989-90 school year?

YES..... 1 PLEASE RECORD INFORMATION CONCERNING THESE AWARDS IN THE BOX BELOW, AND CONTINUE WITH ITEM 4.

NO..... 2 PLEASE SKIP TO ITEM 4

Award Name	Source CIRCLE ONE CODE BELOW Federal State Institutional Other	1989-90 Final Amount	Award Period
1. _____	1 2 3 4	$_____.00	__/__ to __/__
2. _____	1 2 3 4	$_____.00	__/__ to __/__
3. _____	1 2 3 4	$_____.00	__/__ to __/__
4. _____	1 2 3 4	$_____.00	__/__ to __/__
5. _____	1 2 3 4	$_____.00	__/__ to __/__

4. **COST OF ATTENDANCE**

A. Is there a Pell Budget for this student for the 1989-90 school year?

YES..... 1 PLEASE COMPLETE THE BUDGET INFORMATION BELOW AND CONTINUE WITH ITEM B.

NO..... 2 PLEASE SKIP TO ITEM B.

Pell Budget Period ____/____ to ____/____

1. Tuition and fees $_____.00
2. Allowances for room, board, books, supplies, transportation, and miscellaneous personal expenses $_____.00
3. Allowance for child care $_____.00
4. Allowance for Handicapped student $_____.00
5. Student Aid Index (SAI) $_____.00

B. Is there an Institutional Budget for this student for the 1989-90 school year?

YES..... 1 PLEASE COMPLETE THE BUDGET INFORMATION BELOW

NO..... 2

Institutional Budget Period ____/____ to ____/____

1. Tuition and fees $_____.00
2. Books and supplies $_____.00
3. Room and board $_____.00
4. Transportation $_____.00
5. Miscellaneous and personal expenses $_____.00
6. Dependent care $_____.00
7. Handicapped $_____.00
8. Expected Family Contribution (EFC or FC) Total $_____.00

OR

8a. Parental Contribution $_____.00

8b. Student's Contribution from Income $_____.00

8c. Student's Contribution from Assets $_____.00

Figure 5.1 (con't)

5.4 Data collection

The data collection task for the Records Update ran concurrently with the records data collection for February and June 1990 sampled students. The updating and nonfall collection tasks began in August 1990 and were completed by January 1991. The records updates and the nonfall Record Abstracts were completed either by trained data collectors or by institution staff.

5.4.1 The Project Update

Through the NPSAS Project Update, all participating institutions received a reminder of the Records Update Task in July 1990. In addition to providing a brief summary of the project, the information sheet reminded institution coordinators of the upcoming update and nonfall data collection activity and indicated that additional materials were forthcoming.

5.4.2 Information Packet

Those participating institutions not meeting the conditions for data collection by Westat field staff (about 600 institutions) and those institutions that had previously indicated a preference for having the data collection done by their own staff were mailed a Records Update Task/Nonfall Data Collection Information packet. The packet included a letter from NCES which explained the importance of the two activities and identified the time frame in which we hoped to receive completed updates (no later than October 15, 1991). Also included in the package was an example of the Records Update Form and instructions for completing the forms and returning them to Westat.

5.4.3 Telephone Scheduling Contacts

Those institutions using a Westat data collector were telephoned the week after the "NPSAS News" (i.e., project update) was mailed. The purpose of this call was to discuss with the institution coordinator the extent of the work involved with the update task and nonfall data collection and to schedule a date for on-site data collection. A confirmation letter was mailed to each coordinator. Additionally, the NPSAS receipt control system was updated with this information.

5.4.4 Data Collector Training

Data collection at approximately 500 institutions required a field visit. Fifty seven data collectors and three supervisors were used for this purpose. All field staff hired for the update task and nonfall data collection activities had prior NPSAS experience from the earlier data collection field period.

A one-day training session for the supervisors and a two-day training session for the data collectors were held in Denver, Colorado on August 17, 18 and 19. The training plan consisted of a variety of components. A Data Collector Field Manual based on the manual used for the first field period, including revisions of content and form. A home study package was sent to each data collector a week in advance of training. The last component of the training was the training sessions themselves. Training consisted of interactive lecture, cognitive exercises related to the material presented, and role-play activities. Data collector assignment packages were distributed and reviewed at the end of training. These packages included the assignments scheduled for the first three weeks of the field period.

5.4.5 Forms Mailing

In early September, update forms were sent to approximately 600 institutions. The package of forms also included a copy of the control list, an "other specify" awards list and instructions for completing the forms. Some packets also included Record Abstracts for nonfall sampled students. Record Abstracts were sent only to those schools which had a nonfall sample and had asked to complete the Abstracts for these students. Also included was a cover letter from the director of survey operations describing the contents of the packet, reminding schools of the October 15 completion date, and providing mailing instructions for completed forms.

Beginning in mid-September, calls to institutions were made to confirm receipt of the package. At that time, a few schools requested a data collector, feeling that they could not meet the schedule.

5.4.6 Non-response Follow-up

In mid-October a mailing was sent to all institutions that had not yet returned completed Record Abstracts or Record Update forms. By the end of the month, approximately 87 percent of self-reporting institutions had returned completed data collection forms to Westat.

By the end of October, all institutions that had not yet returned completed abstracts or updates were contacted by phone to discuss their progress with the forms, offer assistance as needed and obtain an expected date of completion. By the end of November, completed forms had been received from about 98 percent of the participating institutions. After NCES review, the non-responding cases were assigned a finalized non-responding status.

5.4.7 Technical Assistance

For technical assistance, institutions were advised to call the toll-free NPSAS Information Line. Technical questions were answered by either person who had trained field staff for the Records Update Task or by the director of survey operations. Technical assistance calls were received from more than 600 institutions. Of these institutions, about 65 percent called two or more times. In total, approximately 1,100 technical assistance calls were received. The calls came both from institutions that were self-reporting and data collectors at non-self-reporting institutions. Questions ranged from the very general (what to do with the update forms) to technical questions about specific data items (should the Stafford award amount listed on the update include processing fees).

5.5 Receipt Control

The main receipt control system was expanded to accommodate a new file and variables for the activities associated with the update. As forms were received from the field, a date was entered into the receipt system to indicate receipt of the materials. The student ID numbers on the forms were then checked against those on the control list to verify that all expected forms had been received. The number of forms received was entered into the receipt system. Discrepancies between the number of forms expected and number actually received were submitted to operation's staff to resolve.

5.5.1 Scan Edit

As forms were received they were scan edited by data preparation staff for updates. If a form had any changes or updates, a code was entered in the receipt control system to indicate that updates had occurred. A different code was used when there were no changes. The update forms were also checked for completeness and readability.

5.5.2 Data Entry

A personal computer laptop system was developed to enter data obtained from the Record Update Task. Twelve laptop stations were used to enter and verify corrected or updated information from the Record Update forms.

The laptop system pulled from the main receipt system those records received and marked with updates. Cases were transmitted to the laptop stations on a school-by-school basis. Data entry staff entered the information from the update sheets directly into the laptop system, in effect creating a new data base. Soft ranges were internal to the system so any out-of-range updates required an override by the entry person.

After entering all update data, the newly created data base file was transmitted to a supervisor station for verification and correction. For the first several weeks of the process, supervisors did 100 percent verification of the data entry staffs' work. This complete verification gradually tapered off to about 10 percent. Once all changes were posted, the entire Record Update file was merged with the Record Abstract file to create the Updated Record Abstract file.

CHAPTER 6: STUDENT SURVEY

6.1 Objectives

The Student Survey collected self-reported data from students concerning education, costs, financial aid and other sources of support, employment, and demographics. For first-time first-year students questions were asked about student values, educational participation, and attitudes.

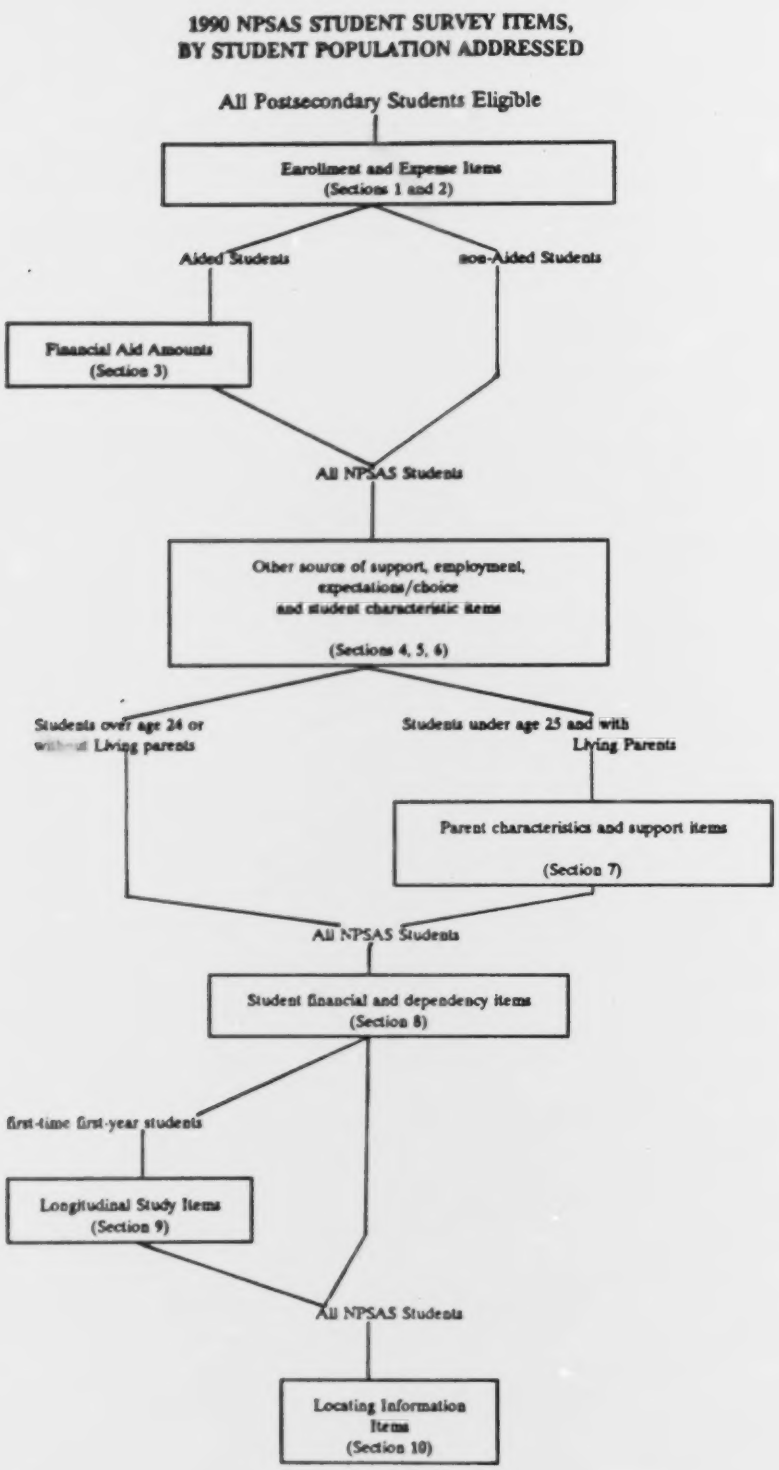
For unaided, independent students, the survey was critical for obtaining information on their financial characteristics. For these students institutional financial aid records were not available, and the student was the primary source of information about sources of funding for education and education-related expenses.

6.2 Instrument Design

A field test conducted during the summer of 1989 tested the application of the 1987 instrument in a CATI environment. Review of the results of the field test and consultation with the Steering Committee produced recommendations which were incorporated into the 1990 Full Scale Student Survey. These recommendations included the addition and deletion of data elements, and proposed changes in interviewer training procedures, design of the CATI system, and number of pre-loaded data elements. The flow of the instrument is shown in Figure 6.1.

During the field test, interviewers found that questions referring to parents were difficult for students to answer especially in cases where parents were separated of divorced. Therefore, a specific data element "parent status" was added to indicate whether both parents or one parent, if any, would be designated as the "referent" parent for further questions. The "referent" parent was defined as: "the parent with whom the student lived the most over the period from April 1988 through March 1989; or the parent who provided the most support from April 1988 through March 1989; or provided the most support in the most recent year in which the student was supported by a parent."

Figure 6.1 Student Population



Wherever possible, data elements obtained during the records data collection at the sampled institution were used to "pre-load" the student data files. Students were then asked to verify information such as date of birth, receipt of federal financial aid, and other demographic data. This reduced the potential for data discrepancies and provided the opportunity to complete missing information in the student data file.

The item wordings are available as part of the electronic codebook. The ten major sections of the CATI instrument are described below:

Section 1: School Enrollment including name of institution and enrollment period by term or program of study during the NPSAS year, July 1, 1989 to June 30, 1990.

Section 2: Enrollment Status and Costs including respondent costs (e.g., tuition) and other expenses and the nature of the respondent's enrollment (e.g., full-time status, credit load).

Section 3: Financial Aid including grants, loans, and other types of aid.

Section 4: Other Sources of Support for educational expenses excluding financial aid such as grants, loans, and fellowships.

Section 5: Employment which obtained information about respondent's employment during the year including dates of employment, earnings, and type of job.

Section 6: Demographics and Plans collected data regarding race, education decisions, plans, and community service.

Section 7: Parent Characteristics obtained information about the referent parent, socioeconomic data, and the amount of financial support provided by parents.

Section 8: Dependency and Financial Status including questions to determine the respondent's dependency status, assets, income, and taxes.

Section 9: Longitudinal Baseline Data was asked only of first-time first-year students and obtained data regarding student values and educational participation.

Section 10: Locating Information included questions to assist in locating students at a future time.

6.3 Staff Organization

The Telephone Research Center (TRC) operations were managed by a senior survey manager who was supported by several assistant supervisors. These supervisors were responsible for

managing, recruiting, and training the interviewing staff, organizing tracing operations, and reviewing refusals.

6.4 Interviewer Training

Interviewer training was conducted from February 1990 through June 1990. During that period, 570 interviewers were trained. Interviewer training consisted of three individual 8-hour sessions, and interviewers were trained in groups of 30 to 35 each. The sessions were led by a key member of the project staff with the assistance of other project staff and staff from NCES. In addition, two to four supervisors were assigned to each group to oversee the training activities of the interviewers.

The basic approach of the training program was to use scripts of the CATI questionnaire. Scripts were presented by a project staff member using an "interactive" lecture technique that encouraged the participation of the trainees. After the interactive lectures, role-plays were done in pairs. Each pair of trainees completed at least four scripted role-plays. With the first script one member of the pair played the role of the interviewer while the other was respondent.

Training materials included the Interviewer's Manual, training scripts, and role-plays. The Interviewer Manual provided explanations and instructions for each CATI screen and a description of valid responses and skip patterns. A glossary of terms related to the subject matter of the study was also included.

6.5 Data Collection

Data collection started in March 1990 and continued through September 1990. The sample consisted of 72,047 students with 54,633 undergraduates, 11,480 graduates, and 5,934 first-professional students. Figures 6.2 and 6.3 show the number of students sampled, eligible for NPSAS, and included in the final analysis files described in Chapter 8, by key characteristics.

Each student received a mailing consisting of a letter from the U.S. Department of Education, a project information summary on the National Postsecondary Student Aid Study, and a student guide to federal financial aid programs approximately one month prior to the first contact by an interviewer. After the initial mailing, an operation was setup to process postmaster return information.

Over the course of the data collection, postmaster returns and remails represented ten percent of all sampled students. A locator file was created containing a record for each sampled student with all known address information collected from the Record Abstract. The file also recorded all remailing dates and addresses and aided in the production of new mailing labels when needed.

Generally postmaster returns were remailed to a forwarding address provided by the postal service. Those with no available forwarding addresses were logged, and the data file was checked to determine if additional locating information was available. If additional addresses were found, a second package was mailed to the student.

Attempts to mail the advance package to students continued until no additional address information was available. The mailing operation also processed requests for packages received from the 800-number toll free operations and the CATI system.

Figure 6.2

Number of Students Sampled, Number Eligible, and Number Included the Analysis File for NPSAS:90

Institutional Level and Control	Sampled	Eligible	Included in Analysis File
TOTAL	72,047	68,929	61,120
Public, < 2-year	1,364	1,165	948
Public, 2-year	7,551	7,180	5,377
Public, 4-year, no Ph.D.	10,054	9,844	8,672
Public, Ph.D. granting	14,326	13,970	12,520
Private, < 2-year	664	562	508
Private, 2-year	1,798	1,720	1,561
Private, 4-year, no Ph.D.	11,723	11,292	10,332
Private, Ph.D. granting	16,093	15,405	13,889
Proprietary, < 2-year	5,511	5,081	4,735
Proprietary, 2-year +	2,963	2,710	2,578

Source: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1990.

Figure 6.3

Number of Eligible Students Included in NPSAS:90 Analysis File

Student Characteristics Education Level	Eligible	Included in Analysis File
Undergraduate	43,868	40,435
Graduate	11,126	9,178
First-professional	5,535	5,096
Unknown	8,400	6,411
Gender		
Male	29,845	26,882
Female	35,893	32,536
Unknown	3,191	1,702
Race		
American Indian	328	295
Asian/Pacific Islanders	2,600	2,210
Black	4,416	3,980
Hispanic	3,730	3,287
White	40,355	36,159
Unknown	15,189	15,189

Note: Based on institution records information, prior to telephone interview.

Source: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1990.

Telephone Center Operations

The survey operations manager was responsible for monitoring the flow of work in the TRC. On a weekly basis the manager reviewed all scheduler queues and determined the most appropriate allocation of labor required to complete the available cases. Considerations in scheduling staff included distribution of cases by time zone, appointments scheduled with respondents, and volume of refusal conversion and tracing work.

Throughout the course of the data collection approximately 536,681 calls were placed by TRC interviewers. Figure 6.4 shows that most telephone contacts were made on weekdays, with 33% of the completes obtained between 6 and 9 p.m.

Figure 6.4 Distribution of completed cases by day and time.

Day/Time (EST)	Percent of Cases Completed
Weekends	29%
9 am - noon	6%
1 pm - 5 pm	17%
6 pm - 9 pm	5%
Other times	1%
Weekdays	71%
9 am - noon	10%
1 pm - 5 pm	23%
6 pm - 9 pm	33%
Other times	5%
Total	100%

Contacting the Students

After locating the respondent at a telephone number, the interviewers worked to get the respondent to complete the interview. If a student was unable to complete the interview at the time of first contact, the interviewer attempted to schedule an appointment for a later time. If the student was not available to schedule an appointment, the interviewer asked the person who answered the telephone to determine the best time to try to call back.

When the interviewer made initial contact with the students, generally students were willing to continue with the interview. In 5% of the cases, the student requested re-mailing of the advance letter before continuing with the interview.

Result Codes

All attempts to contact the students were recorded and result codes were assigned to monitor the data collection progress. In some cases the code was automatically assigned by the CATI system and in other cases was directly entered by the interviewer.
The codes used are as follows:

CO	Completed interview.
DD	Deceased.
IC	Out of the Country.
IH	Ineligible - High school student.
IN	Ineligible - Not a student.
IR	Ineligible - Not at institution.
IS	Ineligible - Sampling problem.
LE	Final Language - Other than English - Unable to complete the interview in English.
LP	Final Language Problem - Unable to complete the interview due to a hearing or speech problem.
MC	Maximum Calls - The calling algorithm had been fulfilled and it included some contact with the household.
NW	Only located number was non-working.
NT	Unable to contact the student.
PC	Partial Complete - Interview completed through Section 3.
RB	Final Refusal - Someone other than the student refused, e.g., the parent.
RF	Final Refusal - Source of refusal not certain.
RS	Final Refusal - The student refused to be interviewed or broke off during the interview and refused to continue.
TF	Not Located.

Refusals

When an initial refusal was recorded by an interviewer, the case was routed to a "review" queue for refusal conversion. The case was scheduled for re-contact in 14 days. Interviewers were trained in refusal conversion procedures including ways to encourage participation in the survey. Major reasons for refusals included:

- Hostile about completing any survey (20%);
- Negative feelings about this survey (14%);
- Felt the survey was too long (10%);
- Felt the survey was too personal (6%); and
- Unwilling to state a reason (50%).

6.6 Locating Students

The CATI data base contained all known telephone numbers and address information for a student. This information included the student's local and permanent address and telephone numbers, and parent locating data. If all available telephone numbers and addresses were exhausted and no additional locating information was obtained through these contacts, directory assistance calls were made using the available information.

Directory assistance calls were made first for the city of the permanent address. If directory assistance had no listings, interviewers next requested numbers for any family with the last name in the town. If this was not successful, calls were made to the city of the local address.

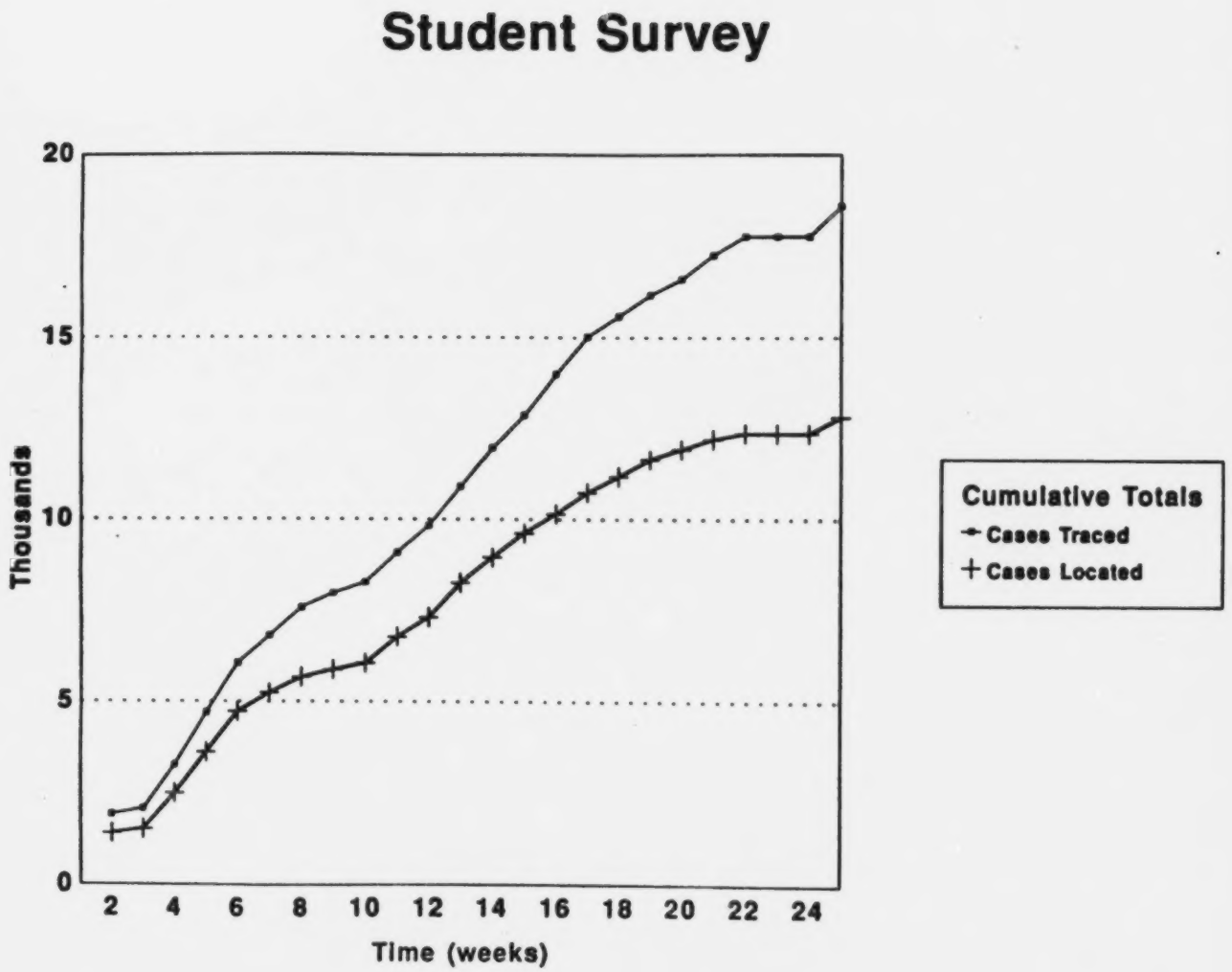
Students were generally willing to provide whatever locating information was requested during the survey for further contact with parents or long term re-contact at a later time. A long term locating address was provided by students in over 86% of the cases where information was requested.

Figure 6.5 shows the number of cases traced and located throughout the data collection period. Of the cases located, 80% resulted in a completed interview.

6.7 Language Problems

Interviewers could classify a case as a language problem if the respondent did not speak English or had a speech or hearing difficulties. All cases classified as language problems were reviewed by a supervisor and re-called to determine if other household members could translate to the student. Spanish speaking cases were referred to a Spanish-speaking interviewer.

Figure 6.5


Student Survey
Thousands
20
15
10
5
0
2
4
6
8
10
12
14
16
18
20
22
24
Time (weeks)
Cumulative Totals
Cases Traced
Cases Located

6.8 Toll-Free 800 Number

The toll-free number was used if the student had no home telephone but could be reached through the mail or friends and relatives. The student could call the number to complete the interview. Also, the number was used as a contact point for students and parents who wished to verify the legitimacy of the study and its purpose. Many calls received by the 800 number were questions regarding financial aid programs. These calls were referred to the Department of Education's information number for student financial aid programs.

6.9 Quality Control

Westat's TRCs are equipped with silent monitoring equipment that enables supervisors to listen to telephone interviews. Approximately 10 percent of interviewers' calls were monitored in this way in order to verify that the interviewer was talking to the respondent and completing the interview properly. An interviewer monitoring report was completed for each monitoring session. Supervisors discussed feedback from the monitoring sessions with each interviewer and were available to respond to questions throughout the collection period.

6.10 Reports

Weekly reports were produced to track the daily status of work by main result, daily callbacks, weekly productivity of individual interviewers, and response rates. Special reports were generated as requested by project or NCES staff.

6.11 Results of Data Collection

Of the 72,047 sampled cases, 66,718 were eligible for the CATI interview. Among these cases, 51,430 completed interviews were obtained. Weighted and unweighted response rates are shown in Figure 6.6.

Figure 6.7 shows the percentage of cases completed by month. Average administration time for completed cases was 38.5 minutes. Times by section are given in Figure 6.8.

Figure 6.6.--Student CATI response rates, by student characteristics for NPSAS: 90

	Eligible	Participating	Unweighted student response rate	Weighted student response rate
All students	**66,718**	**51,430**	**77%**	**76%**
Institution type:				
Less-than-2-year	6,480	3,872	60%	55%
2-year	10,538	7,765	74%	72%
4-year, non-PhD granting	20,955	16,719	80%	79%
Phd granting	28,745	23,074	80%	80%
Institution control:				
Public	31,029	24,529	79%	76%
Private, not-for-profit	28,178	22,170	79%	78%
Private, for-profit	7,511	4,731	63%	59%
Level and Control:				
Public, less-than-2-year	1,082	790	73%	63%
Public, 2-year	6,723	4,912	73%	72%
Public, 4-year, non PhD granting	9,622	7,720	80%	80%
Public, PhD granting	13,602	11,107	82%	81%
Private, Less-than-2year	511	295	58%	52%
Private, 2-year	1,662	1,280	77%	77%
Private, 4-year, non Phd granting	11,013	8,729	79%	78%
Private, PhD granting	14,992	11,866	79%	79%
Private, for-profit, less-than-2-year	4,887	2,787	57%	52%
Private, for-profit, 2-year or more	2,624	1,944	74%	74%
Student Level (as sampled):				
Undergraduate	50,489	38,321	76%	75%
Graduate	10,519	8,471	81%	81%
First-professional	5,710	4,638	81%	81%
Aid and dependency status (based on Record Abstract)				
Aided dependent	12,387	10,097	82%	82%
Aided independent	13,715	10,085	74%	72%
Aided unknown	9,450	7,969	84%	85%
Not aided (23 or younger)	14,063	11,336	81%	80%
Not aided (24 or older)	15,304	11,895	78%	76%
Not aided (age unknown)	1,799	48	3%	2%
Gender:				
Male	28,901	23,145	80%	79%
Female	34,892	27,917	80%	80%
Unknown	2,925	368	13%	12%

Figure 6.6.--Student CATI response rates, by student characteristics for NPSAS: 90--continued

	Eligible	Participating	Unweighted student response rate	Weighted student response rate
Local Residence:				
Campus Housing	11,542	9,673	84%	83%
Off campus (not with parents)	18,375	14,290	78%	76%
With Parents	6,339	4,814	76%	79%
With relatives (not spouse)	199	126	63%	67%
Not specified	30,263	22,527	74%	74%
Student Level (based on Record Abstract):				
Undergraduate: Freshmen (1st)	17,986	13,259	74%	74%
Undergraduate: Sophomore (2nd)	9,420	7,392	78%	77%
Undergraduate: Junior (3rd)	6,960	5,606	81%	80%
Undergraduate: Senior (4th/5th)	8,421	6,790	81%	80%
Graduate: Masters	7,192	5,790	81%	80%
Graduate: Doctoral	1,211	999	82%	82%
Graduate: Unclassified	2,260	1,786	79%	79%
First-professional	5,457	4,436	81%	82%
Unknown	7,811	5,372	69%	69%
Marital status:				
Not married	35,757	27,899	78%	78%
Married	9,329	7,395	79%	78%
Separated	570	324	57%	59%
Unknown	21,062	15,812	75%	73%
Race/ethnicity:				
American Indian	314	241	77%	71%
Asian/Pacific Islanders	2,428	1,881	77%	72%
Black, non-Hispanic	4,307	2,976	69%	66%
Hispanic	3,614	1,930	53%	58%
White, non-Hispanic	39,320	32,224	82%	80%
Unknown	16,745	12,178	73%	72%

Note: Some eligible students identified by institutions were not eligible for CATI because of the following reasons: students deceased, out-of-country, never attended, still enrolled in high school, or not eligible for NPSAS. Chapter 9 describes weighting and variance estimation procedures.

Source: U.S. Department of Education, National Center for Education Statistics, 1990 National Postsecondary Student Aid Study

Figure 6.7 Percentage of completed cases by data collection month

Month of Data Collection	Percentage of Cases Completed
March	15%
April	27%
May	16%
June	19%
July	14%
August	7%
September	2%
Total	100%

Figure 6.8 CATI administration time, by section

Section	Average Administration time in Minutes
1	2.6
2	9.4
3	5.1
4	3.8
5	6.3
6	5.4
7	2.4
8	2.9
9	11.2
10	1.5

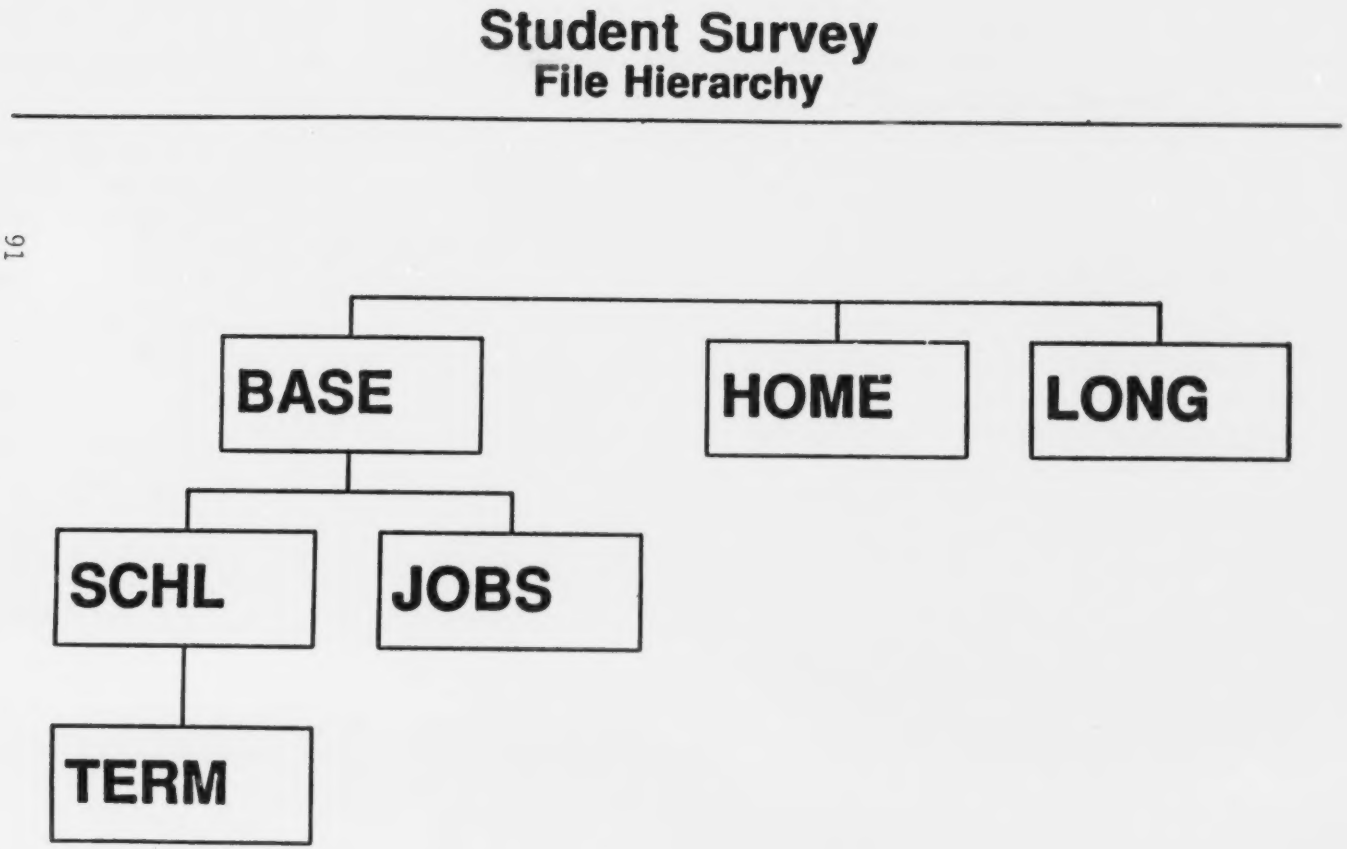
6.12 Data Preparation

The data base for the Student Survey contains six files as shown in Figure 6.9. Case level data is included in the BASE file for all sampled students. For interviewed students additional case level information is contained in the HOME file. Information from Section 9 - Longitudinal is in the LONG file. One record exists for each student who completed section 9 of the instrument. The JOBS file contains employment data for each student with one record per reported job. The SCHL and TERM files include attendance and cost data for schools and enrollment periods. Records exist for each school and term reported by respondents. The variable NPSASID links all files.

Basic range and skip pattern editing was performed by the CATI system. Additional editing and coding procedures are discussed in Chapter 8.

Preliminary data files were delivered to NCES in December 1990, and final data tapes were delivered in June 1991. The final student data files were also included on CD-ROM delivered in October 1991.

Figure 6.9


Student Survey
File Hierarchy
BASE
HOME
LONG
SCHL
JOBS
TERM

CHAPTER 7: PARENT SURVEY

7.1 Objectives

The Parent Survey was a supplemental data collection designed to obtain information from the parents of primarily unaided, dependent students. The survey collected data regarding the costs and financing of postsecondary education, and family financial, educational, and employment characteristics.

7.2 Instrument Design

The 1987 instruments were used as the starting point for instrument design, and were modified for CATI administration. Questionnaire development was a collaborative effort among NCES, the Steering Committee, and Westat.

Question sequences throughout the Student Survey and during the introduction to the Parent Survey determined the appropriate respondent or "referent" parent. In cases of divorce or separation of the parents, the "referent" parent was defined: "the parent with whom the student lived the most over the period from April 1988 through March 1989. If the student did not reside with either parent the referent parent was the parent who provided the most support from April 1988 through March 1989. If neither parent provided support from April 1988 through March 1989, the referent parent was the parent who provided the most support in the most recent year in which the student was supported by a parent.

The parent interview contained five major sections as follows:

- Parental support--including information about parental contributions and loans, and sources and amounts of funding provided;
- Dependents--obtained data regarding number of dependents, level in school, and amount paid for tuition;
- Employment and financial condition--collected information about occupation, income, assets, and taxes;
- Demographics--including race, age, education, and sources of support for parental education; and
- Student's education--including familiarity with financial aid programs, reasons for not applying for aid, and reasons why the student selected the sampled institution.

7.3 Staff Organization

A survey manager was assigned to the survey to oversee telephone interviewing operations. Several supervisors assisted the survey manager. The tracing supervisor was responsible for organizing tracing activities including supervising the tracing interviewers. A second supervisor was responsible for reviewing those cases that were placed in the "problem" and the "initial refusal" categories in order to determine the next appropriate course of action.

7.4 Interviewer Training

From October through November 1990, groups of interviewers were trained in eight hour sessions. For interviewers who were new to Westat, an extra twelve hours of training time was added in order to allow for instruction in general interviewing, CATI administration and a more detailed discussion of the concepts in the questionnaire. Each of the sessions was conducted by either the survey manager or the survey director with the assistance of a supervisor. During the Parent Survey, 122 interviewers were trained.

Components of Training

Efforts were made to standardize each project training session. Thorough and explicit agendas were developed for both the NPSAS experienced trainees and the interviewers new to Westat to indicate the order of events and materials to be used. All presentations were made using scripts, so that the content and form of the information communicated could be controlled.

Our approach to the project training was based upon our prior experience with large CATI training sessions and included two basic training techniques: interactive lectures and dyad role-playing.

The basic approach of the training was to use scripts of the questionnaires to teach the concepts. The first script was presented by the trainer using the interactive lecture technique. This script presented the basic concepts of the questionnaire using an easy example. An overhead screen attached to a portable computer and operated by the training assistant was used to demonstrate how answers should be recorded in the questionnaires. Trainees were called upon to take the role of interviewer and the trainer acted as respondent. While acting as respondent, the trainer emphasized various points and provided specific instruction. Each trainee was assigned to a portable computer and recorded their answers into a training version of the CATI. They were instructed to check their own recording against the correct recording on the overhead projector screen.

More complicated examples and instructions were then presented to the group. Again the trainees took the role of interviewer, and the trainer, using a script, played respondent. The trainers' scripts contained instructions to the leader to stop and have the group read certain sections of the "Question by Question" Specifications in the Interviewer's Manual. All information that the trainer presented was contained in the scripts.

The role-plays were done in pairs or dyads. Each pair completed six scripted interviews. With the first script one member of the pair played the interviewer while the other was respondent. Then with the next script, the trainees changed roles. All role-plays were done on CATI to provide practice with the actual CATI program and allow the trainee to get a feel for the flow of the interview without being interrupted.

Training Materials

Interview training was organized around three documents: The Interviewer Manual, training scripts and role-plays.

Each interviewer was given an Interviewer Training Manual. This manual provided explanations and instructions about all aspects of the interviewer's data collection activity. The first part of the manual discussed the general specifications of the survey and the rules and techniques of interviewing. This was followed by sections specific to the screening process for the Parent Survey and the specifications for each question in the screening section of the interview. Instructions on the extended interview covered the majority of this manual as specifications for every question were provided. A glossary of terms related to the subject matter of the study was also included. The Interviewer's Manual was in a loose-leaf notebook so pages could be changed easily.

The training for the NPSAS experienced interviewers contained two scripts: one covering the interview and one discussing the screens used when contacting the parent. The training for the NPSAS inexperienced interviewers contained five scripts: four of which discussed in detail various scenarios which could occur and screens which could appear during an interview and one discussing the parent contact screens. In addition role-plays were used in each training group to reinforce the concepts discussed in the training. These were scripted interviews, with one interviewer acting as the respondent and the other as the interviewer.

7.5 Data Collection

The main data collection effort for the Parent Survey began in October 1990 and continued until January 1991. The sample included parents or guardian of several subsets of the NPSAS student

population. To be eligible for the parent survey, the student had to have been an undergraduate and less than 30 years old and have completed the telpehone interview. There were approximately 30,640 undergraduates, less than 30 years old who completed the telephone interview. Further, students who were independent, aided, and 25 years or older (unless the student was a first-time student) were excluded from the parent survey. This reduced the number eligible for the parent survey to about 25,225. Due to budget constraints, however, parents of 22,366 student were initially sampled. Further requirements to be included in the parent survey--listed in hierarchial order were that the student was:

- enrolled in school for the first time during the student survey (all--approximately 9,900); or
- listed on school records as a dependent of the parents and receiving federal financial aid but for whom there was no information about the parent contribution to the students education (all--approximately 3,600); or
- non-first-time, first-year students listed as a dependent of the parent and unaided in the school record (a subsample of about 7,100 out of 8,000); or
- listed as independent and aided and less than 25 years old (a subsample of about 1,700 out of 3,100)

Among the initially sampled parents, about 21,900 were subsequently determined eligible for the parent survey. Because the parent survey was intended primarily as a supplement to information collected from the institution and the student telephone interviews, separate parent weights were not developed. It would be possible, however, to develop one or more parent weights, adjusting for parent non-response for the groups mentioned above using ratio adjustment procedures and based on student sampling strata. Interpretation of estimates based on such adjustments, especially for parents of other-than-first-year students, may be difficult.

Initially interviewers were instructed to contact the parent who was designated as the "referent" parent in the Student Survey; however, either parent was an acceptable respondent for the Parent Survey if participation depended on the accessibility of the parent.

Telephone Center Operations

The survey was conducted from Westat's Telephone Research Centers (TRCs). The interviewing staff consisted of a combination of experienced interviewers who worked on the Student Survey and interviewers new to Westat. The Telephone Center supervisors were all experienced members of the TRC staff and many had experience supervising the Student Survey.

Most of the telephone calls were made in the evenings and on weekends, as shown in Figure 7.1, since this was the most appropriate time to contact working parents. However, there were many

Contacting the Parents/Guardian

When an address for the parent was available, either from the student's Abstract or the Student Survey, an advanced letter from the U.S. Department of Education preceded any telephone contacting.

When a case was ready for contact, the interviewer selected the most appropriate telephone number for the parent from those available. The interviewers were instructed to start with the telephone number provided for the parent by the student during the Student Survey. As the telephone numbers were called and eliminated, the interviewers selected the next best available telephone number. Interviewers called referral numbers and all other available numbers listed in order to obtain completed interviews. The case went into tracing when all numbers were finalized and none resulted in a completed interview.

When contact was made with the parent or guardian, the interviewer attempted to conduct the interview. However, this was not always possible and callbacks with specific or general appointments were made with the parents as appropriate.

Result Codes

Each attempt to contact the parent produced a result code. This code was in some cases automatically assigned by CATI based on the information that was entered by the interviewer. In other instances the interviewer had to select and assign the appropriate code. The result codes used on the Parent Survey were similar to those used during the Student Survey.

The codes used are as follows:

CO Complete - The interview was completed with the parent.

DD Deceased - Parents are deceased.

IC Out of the Country

LE Final Language - Other than English - The parent was unable to complete the interview in English.

LP Final Language Problem - The parent could not complete the interview due to a hearing or speech problem.

MC Maximum Calls - The calling algorithm had been fulfilled and it included some contact with the household.

NL Not at number - The respondent does not know parent, does not know how to reach the parent or the parent could not be reached at a particular number.

PC Partial Complete - The interview was completed through Section 3 of the questionnaire.

RB Final Refusal - Someone other than the parent refused, for example the student or other household member.

RS Final Refusal - The parent refused to be interviewed or broke off during the interview and refused to continue.

TF Not located

Refusals

All cases that were coded initial refusals by the interviewer were reviewed by a TRC supervisor. This review resulted in the reassignment of the case to a refusal converter for recall in 14 days.

Refusal conversion training was conducted for interviewers in December, 1990 by experienced supervisors. This training focused on methods to utilize when recontacting parents who refused to participate or students who refused to provide information regarding their parents.

Major reasons for refusals included:

- Invasion of Privacy--Some parents were opposed to answering questions in regard to personal information. In some cases they declined the entire interview. In other cases they were persuaded to answer questions to which they felt comfortable responding. Those questions they felt uncomfortable answering were coded as refused.

- Student was independent or the parent had no role in their education--Parents of students who were independent and not aided financially by the parent very often refused to answer the questions. In some cases they were not even aware that the student had been attending a post-secondary institution.

- Parents of students who were refused financial aid or who did not need aid--Some parents of students who were refused financial aid were bitter about not receiving the aid. Their comments were of the "they did not help us, so why should we answer any questions". Also the parents of the students who were able to meet the financial obligations of the student were of the opinion that they could pay for their child's education and thus had no reason to participate in the survey.

- Confidentiality--Parents were concerned about where the information was going and how it was going to be used. Some refused to answer questions over the phone. Many respondents stated "Send something in writing and I will be happy to answer."

- Anti-government--Some of the parents were opposed to the government spending tax dollars for this study. Perhaps the timing of the study had something to do with this response.

Other examples of the types of refusals received on the parent survey included:

- Not interested

- Too busy - No time

- Interview took too long when the student in the same household was called during the student interview.

- Student in same household refused for parent.

7.6 Locating Parents/Guardians

The purpose of the tracing effort on the Parent Survey was to identify a telephone number at which we could interview the parent or guardian. The tracing sources used included Directory Assistance, telephone books, and Digital Phonebook USA.

The tracing effort for the Parent Survey began in December, 1990. In the Student Survey a series of questions identified the student's referent parent, their relationship to the student, their address and telephone number. All numbers for the parent/guardian provided by the student were loaded into CATI as well as numbers from the student abstract indicating parent/guardian name and telephone number.

A group of experienced tracers who worked on the Student Survey were trained for Parent Tracing. The tracing specifications for the Parent Survey divided the work into categories based on who had been identified as the referent parent or guardian, and marital status of the parent. This distinction was made because some cases required tracers with more experience (for example, the tracer might be contacting the student or a divorced parent who was not identified as the parent to be interviewed). During the training all tracers learned the meaning of each field of information. It was important for them to understand whether the locating information came from the Student Survey, the Abstract or the Parent Survey.

Figure 7.2 shows the total number of cases traced and located over the course of the survey. Tracing efforts were usually successful in locating parents.

7.7 Language Problems

Throughout the course of contacting parents, interviewers would reach respondents with hearing or speech difficulties as well as non-English speaking respondents. Westat trained all interviewers on procedures for handling these situations. The interviewer would attempt to determine if another member of the household was available to speak with. If no other English speaking adult was present the interviewer coded the case a "Language Problem" and wrote a brief explanation identifying the specific hearing, speech or language difficulty. This case was then entered into a queue which required supervisory review through the utility program.

In reviewing the cases the supervisor determined if the case should be called again by an interviewer trained to handle language problems. It also allowed the supervisor the ability to recode and finalize these cases. If, upon recontact with the parent, the interviewer determined that the parent could not participate, the cases were assigned the final result codes of LP (hearing or speech problem) and LE (non-English speaking problem).

7.8 Toll Free 800 Number

The toll free 800 number was used when the interviewer reached a parent but they would not continue with the interview until they called to verify the legitimacy of the survey. If the parent called the toll-free 800 number and wanted to continue with the interview, we had the ability to locate the case ID and have an interviewer continue with the parent. The 800 number was also used for cases where the only way to reach the parent was for them to call Westat. These cases were recalled on a regular basis to see if the parent received the message to call the 800 number and also to obtain another number where the parent could be reached if one was available.

7.9 Quality Control

Monitoring quality control of the data collection effort was the responsibility of the project staff and telephone supervisors. Monitoring the work of interviewers was the primary quality control procedure using extension telephones and video displays linked to the interviewers' CRTs, supervisors monitored about 10 percent of each interviewer's work.
Supervisors randomly select intervals of each interviewer's working time as a monitoring period. The monitoring period includes whatever results the interviewer obtained for a minimum of 15 minutes.

Supervisors completed an interviewer monitoring report each time an interviewer was monitored. This report was discussed with the interviewer to provide feedback when the monitoring period was over. Data items that need to be corrected were recorded. At the end of each week a supervisor recorded in the CATI System the total number of monitoring sheets for each interviewer. The cumulative total appeared on a weekly report.

7.10 Reports

The management reports provided information regarding daily status of work by main result, daily callback report, weekly productivity of individual interviewers, weekly calendar of number of telephone calls made, and weekly response rate report.

7.11 Results of Data Collection

Of the 22,366 cases selected for the parent survey, 494 cases were ineligible. Among eligible cases (21,872), completed interviews were obtained with 16,106 parents. Figure 7.3 shows the average administration time of each section of the survey. Average administration time for the survey was 20.6 minutes. The distribution of completed parent interviews by student institution type and dependency status is shown in Figure 7.4. The weighted response rate for students' parents to the parent CATI was 87 percent (weighted number of students' parents who participated divided by the weighted eligible number of students' parents).

7.12 Data Preparation

The parent survey data base consists of six files as shown in Figure 7.5. This includes a BASE file to define the case, and a separate file for each major section of instrument as follows: COST - Parent Support; SUPP - Dependents; FINC - Employment and Financial Condition; DEMO - Demographics; and ATTI - Student's Education. The records for a case are linked by NPSASID. Each sampled case contains a BASE record. Additional records are present depending upon the sections of the instrument completed by respondents.

The CATI system performed basic range and consistency editing throughout the interview. The editing and coding procedures for the Parent Survey are described in Chapter 8.

Preliminary data files were delivered to NCES in April 1991, and final files in June 1991. Parent data were included on the CD-ROM delivered to NCES in October 1991.

Parent Survey

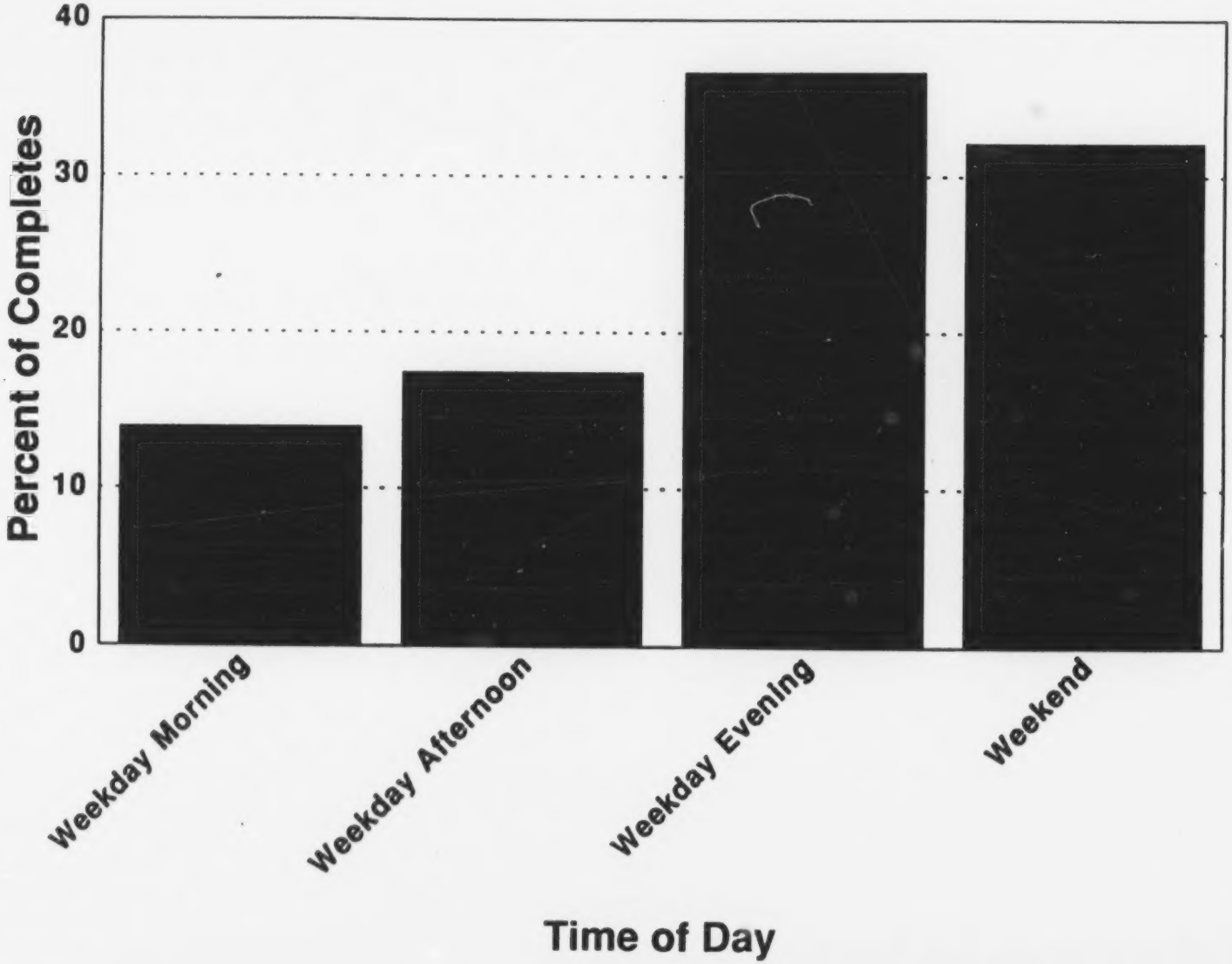


Figure 7.1

Parent Survey

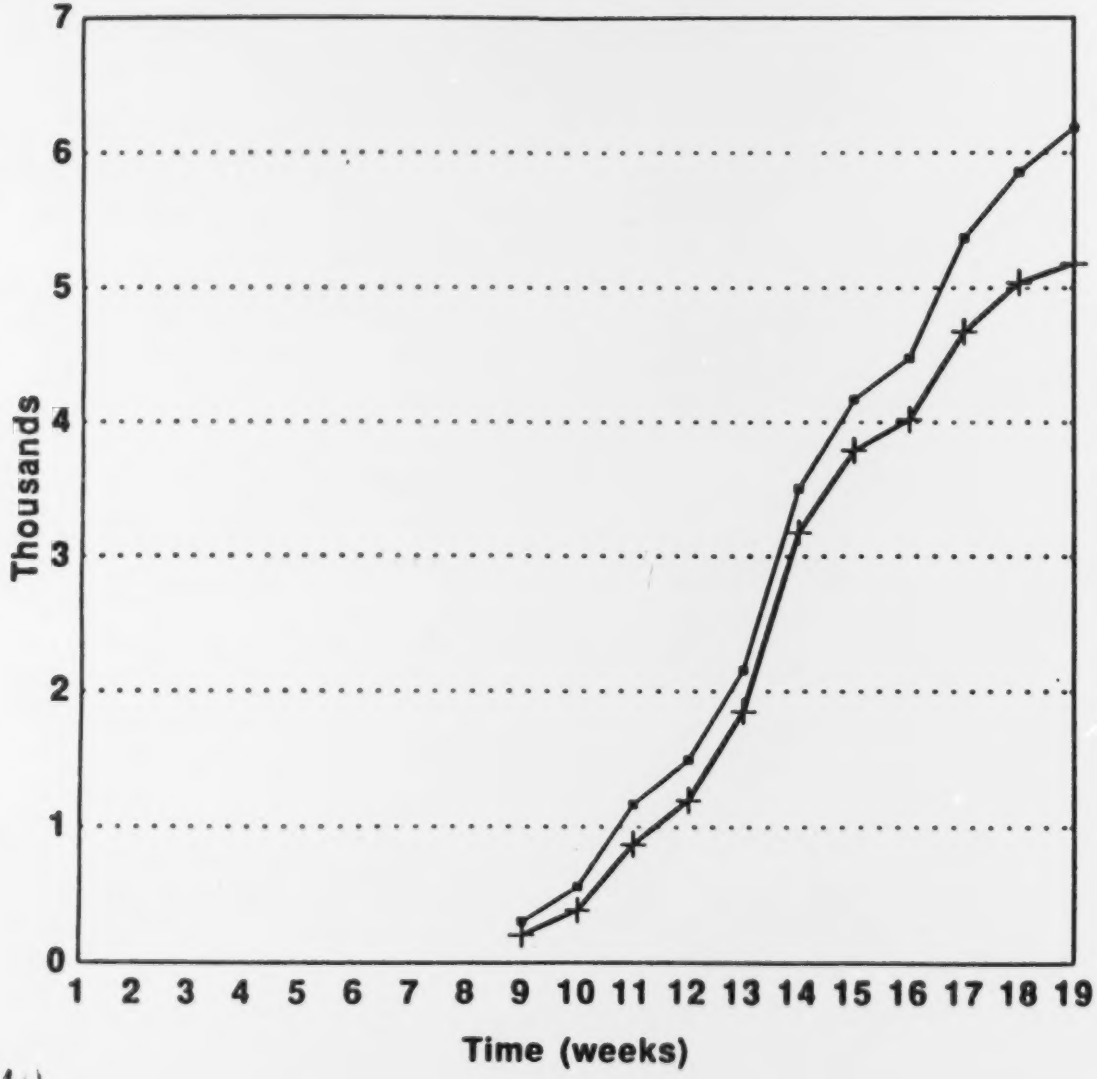


Figure 7.2

Parent Survey

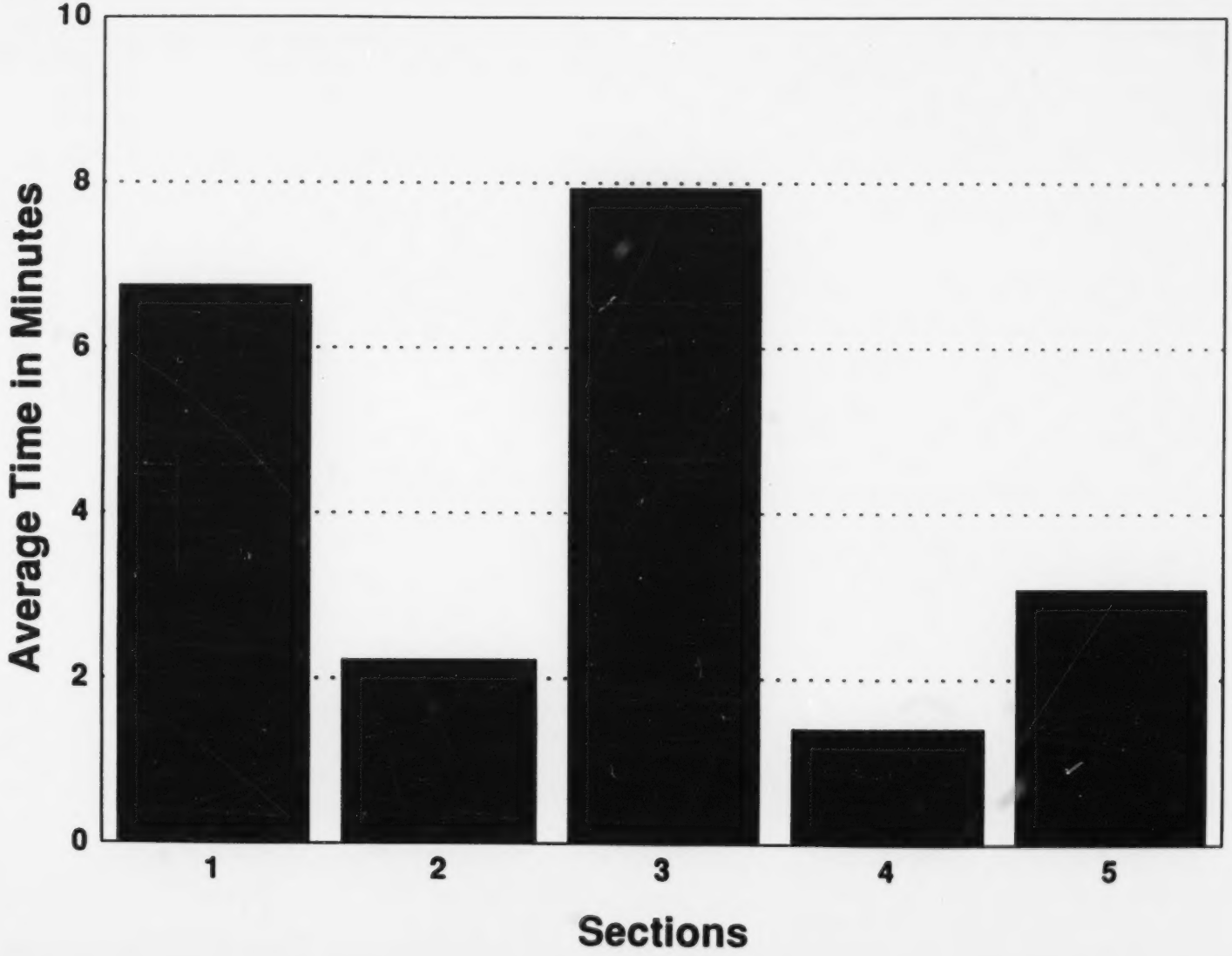


Figure 73

Figure 7.4

Number of Parents Selected for NPSAS:90 Parent Survey, Number Eligible, and Number Participated

Institutional Level and Control	Sampled	Eligible	Participated
Total	22,366	21,872	16,106
Public, < 2-year	357	349	247
Public, 2-year	2,313	2,266	1,609
Public, 4-year, no Ph.D.	3,500	3,450	2,643
Public, Ph.D. granting	4,106	4,032	3,085
Private, < 2-year	157	149	95
Private, 2-year	817	792	622
Private, 4-year, no Ph.D.	3,716	3,634	2,808
Private, Ph.D. granting	4,534	4,375	3,179
Proprietary, < 2-year	1,812	1,783	1,090
Proprietary, 2-year +	1,054	1,042	728

Source: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1990.

Parent Survey
File Hierarchy

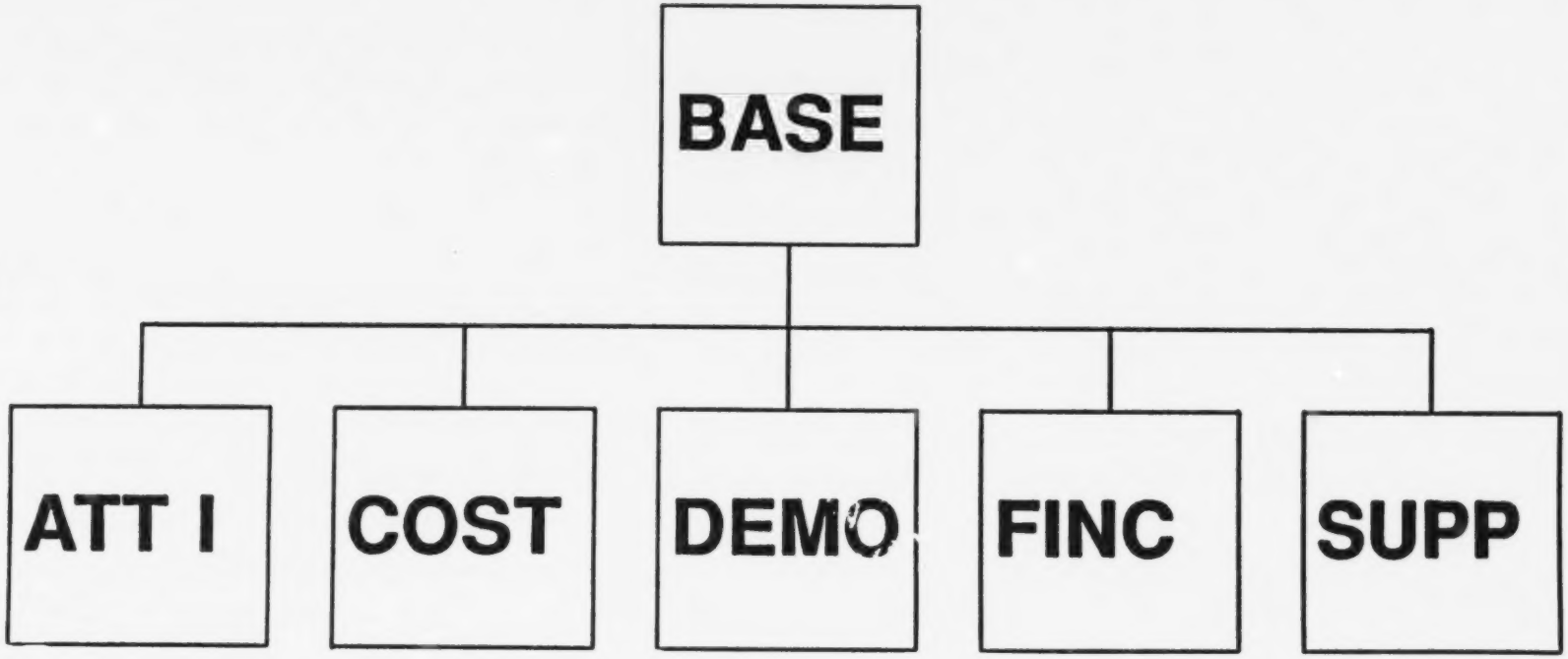


Figure 7.5

CHAPTER 8: FILE CREATION AND DATA ANALYSIS

In this chapter we will review our procedures to edit and code data sets resulting from data collection activities, and describe the development of the NPSAS analysis file. First, we will briefly summarize the main characteristics of each of the NPSAS data files shown in Table 8.1.

The **institutional records and awards file** contains data on school and student characteristics obtained through in-person record abstraction and a records update. Student information includes personal characteristics, such as, gender, race, marital status, age, and financial characteristics, including aid amounts, types and income data. Chapters 4 and 5 describe the collection of records file data.

The **student survey file** contains data obtained through telephone interviews of more than 51,000 enrolled postsecondary students. The data fall into ten categories: school enrollment, costs, financial aid, other sources of support, employment, personal expectations, parental characteristics, dependency status and income, demographic characteristics, and locating information. Chapter 7 described the student data collection.

The **parent survey file** contains the results from over 16,000 telephone interviews with a subset of the student sample. Data in this file include information on parental support, family size and dependents, employment and financial condition, demographic characteristics, and student-related information and family attitudes. Information on the parent data collection has been presented in Chapter 7.

The **analysis file** contains a selected subset of variables from the previous three files. It is intended as the key research data base for NPSAS:90. There are approximately 300 variables in this file (and an equal number of source identifiers, one for each variable--see Section 8.3).

8.1 Editing

The purpose of editing is to achieve a consistent data set which is relatively free of obvious data entry and other logical errors. Westat used both manual and automatic systems for editing the 1990 NPSAS files.

Generally the more complicated the survey instruments and the more numerous the data files, the more editing is required. The amount of editing required is also a function of the type of data collection used. For example, financial aid and registration information data were compiled from school records on site by record abstractors, without the benefit of automatic editing control. These data were subsequently entered into the record abstract data file. No editing occurred on site, instead

it was carried out manually, following receipt at Westat's central office. In contrast, the student telephone data collection used an automated computer editing system (CATI) that guided the interviewers through the intricate skip patterns and signaled when data were out of range or inconsistent with prior responses or other potential problems. Wherever possible, staff built in automatic editing procedures to minimize the more time-consuming task of manual editing.

Table 8.1.--1990 NPSAS data files summary table

File	Respondents	Maximum Number of Observations	Key Elements	Number of Variables
1. Institutional Records Data	Institutions	68,599**	Demographic, Budget and Application Data	431
2. Institutional Awards Data	Institutions	403,477**	Dates, Types and updated amounts of awards	32
3. Student Survey Data	Postsecondary students	69,613**	Demographic, employment and financial aid data	498
4. Parent Survey Data	Parents of postsecondary students	16,106	Parent demographics, sources of support, and attitudes	241
5. Analysis File	Integrated Institution, Student and Parent	61,120	Full-year, updated award amounts, demographics, budget and application data	665

Note: **Institutional records data were collected for 68,599 students. Telephone interview data were collected for 51,430 students and 16,106 parents. However, the institutional awards data file and the some of student survey data files include multiple records per student. The student survey data jobs file contains one record for each job--a total of 69,613 records. The institutional awards data file contains one record for each award received or institutional budget data element. Some students may have up to 32 records in this file, which contains nearly 403,500 records.

8.1.1 Institutional Records Editing

As mentioned above, data from the record abstracts and update files were not edited using an automated system. Prior to release of these files, editing was necessary to remove extreme outliers and other data inconsistencies, especially for the financial aid award amounts. Because award data were collected by award period, our first step was to edit the award period dates to assure that they

did not overlap and that they matched the NPSAS award year (July 1989 - June 1990) as closely as possible. To accomplish this the following edits were performed on the record abstract award data for each award period:

- Correct overlapping award periods by checking dates for other awards collected during the same term and the data for awards across terms for contiguous periods.
- Correct awards with partially missing dates by filling in with dates from other awards from the same term for the same student, or with the modal period derived from the matrix of award periods. Awards with completely missing dates remained missing.
- Correct awards that were outside of the NPSAS year (7/89 - 6/90), by checking for consistency with other awards collected under the same term for the same student and across terms.
- Correct budget variables with missing dates by using information from other budget variables collected for the same term for the same students.

Once award period amounts were established, it was necessary to apply range edits to the award amounts. These included, but were not limited to:

- Awards less than $10 were multiplied by 10 to ensure that the minimum awards were not less than $10.
- For awards with statutory maximum award amounts, the upper award limit was set to at least twice the known maximum since it is possible that a student might have two awards recorded on their file during one award year. For awards with no known maximum, stemleaf plots and box-plots were used to display the distribution of the data. Extreme outlying values were divided by either 10 or 100 to bring them within range.
- For cost variables, minimums and maximums were controlled by statutory limits or by acceptable ranges based on the College Board Cost Book. For example, transportation costs were capped at $7,000 which is twice the highest listed transportation budget in the Cost Book.

8.1.2 Student Survey Editing

Student Survey data were collected using the CATI system which performed range and skip pattern checking as the interview progressed. Following the completion of data collection, editing programs were run on the data files to:

- verify ranges and skip pattern logic;

- check the integrity of the database (to determine that all records for a case were present);
- correct awards with partially missing dates; and
- correct awards that were outside the NPSAS year.

Another major component of the post-interview editing was the incorporation of information collected through interviewer comments into the data set. Due to the complexity of the interview and the complicated skip patterns, it was not uncommon for respondents to provide inconsistent answers, or to remember more information about school attendance, receipt of financial aid, or employment at the end of the interview. When this occurred, the interviewer recorded the new information in a "comments" field.

Approximately 50% of the cases contained interviewer comments which required further processing. Of these cases, 40% resulted in updates to variables. Edit programs were run following all item updates to assure the integrity of all skip patterns and ranges.

8.1.3 Parent Survey Editing

The CATI system controlled range and skip pattern checking for the Parent Survey. The interview was considerably shorter than the Student Survey and most skip pattern sequences were straightforward. At the conclusion of data collection the following edits were performed:

- checking all variable ranges;
- verifying all skip sequences; and
- structural edits to determine the presence of all records for a case.

8.2 Coding

Coding is the assignment of logical values or categories to survey responses which are not automatically coded in the survey, such as open-ended questions. Coding maximizes the use and flexibility of variables on a file, by simplifying unwieldy numbers of responses. For example, coding is required for student occupation and industry classifications which can take on a large number of values. On the NPSAS Student Survey file, students reported over 100 different occupations. To make this data more amenable to analysis, students were sorted into 38 standardized job categories,

such as engineers, service workers, teachers, following a widely-accepted standard list of occupations.[2] By so doing, a file user can more easily analyze student employment characteristics.

Another example of coding is the reporting of "Other Aid" on the record abstract and records update files. Where possible, these unspecified types of aid were reassigned to specific types of aid. For example, an entry in "Other Aid" could have been SLS. Since it is likely that this was meant to be the abbreviation for the Supplemental Loans for Students program--SLS. This aid record would be changed to signify receipt of an SLS. Clearly, there are risks, and loss of information, during coding procedures. In the example above, it is also possible the SLS was in the records as SSL, and meant a Stafford Student Loan. Where the choice is ambiguous, we have placed the other aid into the broadest possible aid category, such as need-based federal loans.

8.2.1 Institutional Records

The following coding changes were made to selected variables on the Institutional Records files:

- Missing codes were standardized, such that, -9 indicates missing and -1 indicates legitimate skips.
- The frequency distributions of the following variables were reviewed and outlying values were set to missing: Q32A1 Q32A2 Q32B1 Q32C2 Q34B Q34C Q34DF Q34DT Q34DTTF Q35CF Q35CT Q35CTTF
- Coding checks were applied to the following variables:

-- Q27 - sample term, if missing, sampling information was used.

-- Q29 - program of study, if missing, program of study is derived as follows. Students who either received Pell or attended a 2-year or less school were classified as undergraduates (Q29=4). For other students, sampling information was used to replace missing.

-- Q30 - undergraduate level was set to legitimate skip for graduate students.

-- Q31 and Q32 - SAT/ACT scores. SAT scores greater than 800 or less than 200, and ACT scores less than zero or greater than 36 were set to missing. Q31C was edited according to whether there were valid data in Q32 after the range edit.

---

[2]Standard occupational and industry codes (SOC and SIC) have been developed by the Census Bureau for use with their employment surveys.

-- Q37 and Q38 - GPA scores. GPA scores in Q37 were checked against Q38 GPA scale for consistency. For example, if Q38 = 1 (GPA rated on a 4-point scale) and Q37A or Q37B were not between 0 to 4, then Q37A or Q37B were set to missing.

-- Q58--Q69. Responses to these items were edited to conform with the availability of budget information in the awards file.

-- Q71--Q76 aid application forms. Q71A--Q71G, and Q72 were edited to conform with the availability of data in Q73-Q76. For cases with completely no valid responses to the questions on an application form, responses were set to -1 (skip). For cases with partially completed application forms, the remaining missing cases were set to -9.

8.2.2 Student CATI File

The coding process had multiple stages. Literal strings from open-ended responses to questions were sorted, and only unique strings were selected for autocoding. The file of all unique strings was matched against prior coded responses. For example, "accountant" would be matched with exact as well as inexact spellings in the coding process. Unique strings not coded by the software were placed in a manual coding and review file. Coders reviewed these strings and entered the appropriate codes. The software applied these codes to all identical strings in the coding files.

Coding schemes were based upon CIP (A Classification of Instructional Programs), SOC (Standard Occupational Classification Manual), SIC (Standard Industrial Classification Manual), and IPEDS. Other specify responses were reviewed by coders and when possible reclassified into existing response categories.

The following variables were coded on the Student Survey files:

- Other specify responses;
- Parent occupation;
- Student industry and occupation for all reported jobs;
- School name;
- Student occupation for future career; and
- Program of studies or major.

8.2.3 Parent CATI File

The coding procedure for the Parent Survey was identical to the process described for the Student data set.

The following variables were coded on the Parent Survey files:

- Other specify responses;
- Parent occupation; and
- Parent industry.

8.3 The NPSAS Analysis File

Our key objective for this file was to provide NPSAS data users and researchers with sufficient and accurate information to meet most of their analytical needs. The resulting NPSAS analysis file contains about 300 transformed variables, concentrating on student demographics and financial aid characteristics. To create this file, we applied several modifications to the edited NPSAS raw data files. The most important of these were:

- Carefully **selecting and deriving a comprehensive**, though abridged, **set of analysis variables**;
- Summing the award data from the updated record abstract and student CATI raw files across award periods to derive **full-year award amounts**;
- Using **secondary data sources**, when available from within NPSAS to fill in missing data and to confirm data which exceeded acceptable limits;
- Establishing **editing ranges** which conform to statutory rules, when they exist, such as the annual limits on Stafford Student Loans which vary by academic level;
- Establishing **imputation rules** for variables which, after primary and other sources are used, contain five percent or more cases missing.

To assist locating variables of interest, we arrayed the approximately 300 variables on this file into 12 categories (see Table 8.2):

- student characteristics, including age, race, and gender;
- school characteristics, including type and control;

- enrollment characteristics, including attendance status, admissions test scores, grade point average, and academic level;
- schooling costs, including both student-reported and allowable costs for computing aid eligibility;
- student financial resources, including scholarships, waivers, loans and employment;
- family financial resources, including contribution and loans;
- aid eligibility characteristics, including dependency status and selected application variables;
- full-year award amounts for specific award types, such as Pell Grants, State grants, private scholarships, or institutional loans;
- award combinations and aid ratios, including receipt of specific packages and indicators of grant to loan ratios;
- source indicators for every variable, which denote the source files and imputations used when creating each variable; and
- other variables, including weights and sampling stratum.

8.3.1 Variable Selection and Development

Westat provided NCES with an initial list of variables for inclusion in the analysis file. Following NCES approval of the final set of variables, Westat drafted an initial set of specifications for these variables.

Table 8.2--NPSAS Analysis Variables

I. Student Characteristics		II. Institution Characteristics		III. Enrollment Variables		IV. Costs	
age	stuind	control	ofcon1	progtyp	satm	totcost	cmtuit
marital	choice	level	ofcon2	ugrdlvl1	sattotal	tuitcost	cmbooks
hsdeg	distance			ugrdlvl2	act	roomcost	cmroom
gender	majrcode/courcode			gpa	attnstat	othrmcst	cmtrans
localres	aveexp			gpacat	compto87	bookcost	cmmisc
ctznshp	actvduty			lengthcl	enlen	othrcost	cmdpndnt
race	veteran			hssperwk	pstsecyr	offcost	cmhandcp
race	disablty			credhrs		injuris	cmefc
applynsh	commserv			attend		pelltuit	cmcosts
acceptat	comserhr			noenroll		pellroom	cmbudget
religion				nosch		pellchil	cstperfc
stuocc				satv		pellhand	

V. Student Resources		VI. Family Resources		VII. Aid Eligibility Variables		VIII. Award Flags	
anycws	cwspernd	parmar	nrefloan	depend	farmval	all award amounts* any	
cawamt	wkinc	refpar	refinc	efc	appform	aid	
scholshp	wkinccal	refcontr	parcontr	sai	forms		
scholamt	evrborw	refloan	parloan	nonfmcst	fedtaxes		
waiver	boramt1	nrefcon	other sources*	pelldiff	untaxinc		
waivamt	boramt2			depinc	famnum		
spsemp	oweamt			indepinc	posted		
spsinc	stsavpln			faminc	othertax		
empiprd	sabonds			famfarm			

IX. Award Amounts				X. Multiple Awards/Aids Ratios		XI. Source/ Imputation Flags	XII. Other Variables
totaid	tfedaid	inneedgr	totgrt	aidsrc	loanpack	all variables	datasrc
othscamt	t4amt1	stateamt	totloan	aidsrc1	grtpct		emstat1-emstat12
pellamt	t4amt2	statneed	totwkst	aidsrc2	loanpct		enstat1-enstat12
seogamt	fedamt1	statnond	tototh r	aidpack	workpct		mnstat1-nmstat12
cwspamt	fedamt2	otheraid	tnfedaid	fedpack	grtloan		ipeds1-ipeds3
perkamt	campamt	teachamt	tfedgrt	pellpack	grtratio		npsasid
staffamt	instamt	resamt	tnfedgrt	stafpack	aidratio		*weight
plusamt	instcws	astamt	tfedln				stratum
slsamt	instneed	fellamt	tnfedln				psu
iclamt	instnond	trnamt	tfedothr				
othfdamt	innondgr	othgramt	tnfedoth				
titivamt							

155

156

Derivation of most of these variables was relatively straightforward. However, there were some decisions and assumptions that had to be made before arriving at the final derivation algorithm for certain variables. Some of these decisions were generally applicable to many variables. For example, certain student characteristics, such as marital status or dependency status can change during a school year. For clarity and consistency, we classified each student by the characteristic as it was at the initial sampled term. The same applied to type and control. Whether or not a student transferred to a different school during the year, only one type and control was indicated on the analysis file. While the full-year award amounts combine awards at all schools attended[3], only the type and control of the sampled school appears on the analysis file.

Student Costs

One of the more difficult sets of variables to create were those related to students' education costs, i.e., their tuition, room and board, and other expenses. While seemingly obvious, there are four different sets of costs collected through NPSAS and these create some problems of selection and definition for the analysis file. The four sets are: student-reported costs, Pell budget, the Congressional Methodology budget, and the institutional budget. The last three are costs which are used to compute aid eligibility.[4]

When creating these cost variables, Westat adhered to two principles:

- First, a researcher should be able to determine approximately a student's basic, unadjusted cost items (tuition and fees, room and board, and miscellaneous) and to differentiate residence status (in or out of state) and local residence (on or off campus).
- Second, a researcher should be able to determine the main budget or budgets used to determine a student's aid eligibility.

To meet these principles, we have provided about two dozen different cost component variables, divided into three groups:

---

[3]For students who attended multiple schools during the year, the full-year award amounts are comprised of awards reported by the sampled school from the updated Record Abstracts file as well as student reported award data for all other school from the Student CATI data.

[4]In need analysis, aid eligibility is equal to the difference between cost of attendance and expected family contribution. Therefore, the level of cost of attendance assigned to an individual student is critical to the level of aid eligibility. In the law, certain allowances are set for various components of cost of attendance (these are the so-called "allowable costs"). For example, in the Congressional Methodology the allowance for room and board costs cannot be lower than $1,500 for a student living with his parents and having no dependents. The Pell budget, in contrast, sets an allowable cost for room, board, books, supplies, transportation, and miscellaneous expenses for this same type of student at a maximum of $1,700.

- student-reported costs (with tuition and fees checked against record abstract data, see Chart 8.1.)

- Pell allowable costs

- Congressional Methodology allowable costs, through which schools certify eligibility for and award campus-based federal aid, Stafford Loan and other institutional support[5].

Award Flags and Source Flags

Often analysts need to know only about the presence or absence of an award, rather than the actual amount of the award. Thus, for each award amount variable included in the derived variable file, we added a corresponding award flag based on the award amount. For four variables (PELL, STAFFORD, PLUS, and SLS), the award flag can have three values: 1 if no aid of that type was received, 2 if some amount less than the maximum was received, and 3 if the maximum award was received. The remaining award flags were all dichotomous, taking the value 1 if a non-zero award was received, and 2 if no award was received.

A source flag was created to correspond with each analysis variable in the analysis file, so that analysts may determine how the analysis file is based on the separate raw data files (the record abstracts and the student and parent surveys). The flag variable denotes which survey form was used to supply the data (or that a combination of forms was used), and whether the data was adjusted (e.g., because an award amount exceeded the federal maximum for a single year). In those cases where data were either imputed from other variables in the analysis variable or acquired from Pell program data, the source flag takes the value "9." A complete list of the codes used for the source flags is provided below.

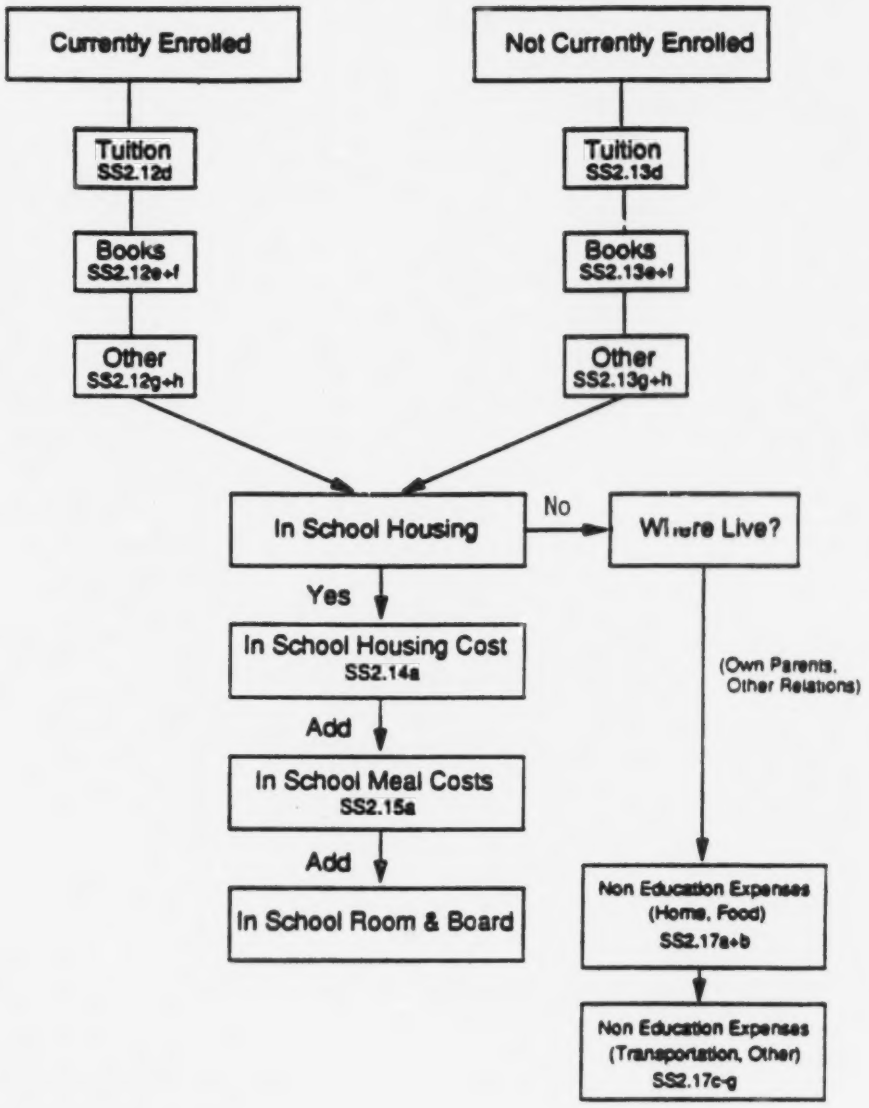
0 = Missing, no imputation perfor.
1 = Record Abstract, not adjusted
2 = Record Abstract, adjusted
3 = Student Survey, not adjusted
4 = Student Survey, adjusted
5 = Parent Survey, not adjusted
6 = Parent Survey, adjusted
7 = More than one source, not adjusted
8 = More than one source, adjusted
9 = Imputed from other variables

---

[5]Though in a small number of cases schools also use their own institutional budgets, most use the CM budget. As a result, we have excluded these institutional budget elements from the analysis file. Rather, we have included a variable which identifies when the institutional budget and the CM budget are the same.

Chart 8.1

Derivation of Student Reported Costs



8.3.2 Full-Year Award Amounts

While the raw data files contain information on costs and financial aid for individual terms, the derived variable file focuses on full-year amounts. This has several advantages:

- Federal regulations (such as award maximums) are often based on a full year;
- The number of variables (or cases) required is greatly reduced, and does not vary from one student to another; and
- The issue of differing term lengths (e.g., semester versus quarter) has less relevance, since there is less variation in the length of an academic year.

Full-year amounts were created in several stages. First, each term was examined for correspondence with the NPSAS year, and all terms falling entirely outside of that time period (July 1989 to June 1990) were excluded. Second, for each individual financial aid award category (or cost category), the sum across all terms was calculated. Third, full-year totals were adjusted to correspond with federal regulations on financial aid awards, especially with regard to award maximums. This was done to prevent misinterpretation of the data. It is possible for a student's financial aid awards to be timed so that awards for two years appear within the single NPSAS year, deceptively making it appear that the student received an overaward. To prevent such misunderstandings, the data were adjusted.

8.3.3 Secondary Data Sources

In the case of financial aid award amounts, the record abstract was treated as the primary source of data. Thus, data from the student and parent surveys were not included in the sums except in the following cases:

- If a student attended more than one school in the NPSAS year, the record abstract provided data only for the sampled school. Data on financial awards at the non-sampled schools were taken from the student survey.
- For 330 cases, no record abstract data were collected. When available, student survey data were used for all terms for these cases.
- For a few award categories, the student or parent surveys were used to check the possibility that an institution's record data may not have included all awards. These categories were graduate assistantships and fellowships (graduate aid was more difficult to identify from institutional records, because the award data were not necessarily kept in a single location), PLUS loans (on the assumption that a school may have been less well informed about parent loans than about direct aid to the student), and other aid (e.g., from businesses or fraternal

organizations) that may have been paid directly to the student without notifying the institution. For these categories, the amounts reported on the record abstracts were compared to those on the student or parent surveys, and the higher amount was used.

Sometimes, combinations of the above rules were followed. For example, graduate assistantships and fellowships recorded on the record abstract were compared to those on the student survey for the sampled school, and then additional assistantships or fellowships were included in the sum if the student attended more than one school in the NPSAS year.

For 4,628 students, there was a need to combine record abstract data on awards and student survey data, because the student survey included additional schools not covered in the record abstract. However, award categories in the student survey did not correspond exactly to those in the record abstract. From the start, the record abstract was planned to be the primary source of financial aid data, so less detail was collected on the type of awards when conducting the student survey. For example, the record abstract collected data separately for teaching assistantships and research assistantships, while the student survey grouped both types of assistantships together. Rather than trying to apportion the student data among teaching and research assistantships, an extra variable (ASTAMT) was created to include all assistantships where the type of assistantship was unknown. (Sometimes the institutional records failed to distinguish the type of assistantship, so amounts reported within ASTAMT may come from either the record abstract or the student survey.) More generally, however, student data were only used when the type of aid could be clearly specified. Thus, in the case of students who attended more than one school during the NPSAS year, there is a possibility that some students' aid will be understated in the full-year amounts because the exact type of aid could not be determined from the student survey.

Another variable (OTHSCAMT) was created that includes all financial aid reported in the student surveys for non-sampled schools attended in the NPSAS year. This variable may be used as a partial check on the importance of the non-sampled schools to the total aid amount (TOTAID). However, the two variables are not fully comparable, and a detailed comparison would require the use of the raw student survey data. There may be cases where OTHSCAMT includes aid not included in TOTAID (because of an inability to determine the proper aid category in the student survey data) or where TOTAID includes additional edits of the award amounts not used for OTHSCAMT. (For example, an award amount may look reasonable when examined alone, yet push the student's total full-year amount above the maximum for a single year when added to the record abstract data. In such a case, the full-year amount was adjusted to correspond with federal regulations, but no adjustment was made to OTHSCAMT.)

8.3.4 Editing Ranges

Three major types of edits were performed on the full-year financial aid award amounts. First, if full-year totals exceeded the one-year maximum awards for a financial aid category, the award amounts were adjusted accordingly. The most likely reason for an award amount exceeding the maximum is that a student received awards for two separate years within the dates set for the NPSAS year. Thus, award amounts were often divided by two if that appeared a likely explanation for the discrepancy. In other cases, the problem appeared to be a problem with the location of the decimal point, and in still other cases the award amount was reset to the maximum. A second edit that was used for some federal aid categories was to compare the student's aid amounts with the student's financial costs, since federal awards are often restricted to not exceed costs. This edit was not always applied, especially if it would have reduced financial aid amounts excessively. A final edit was of award minimums; some aid amounts that appeared excessively low (i.e., below $50 or $100) were adjusted upward based on criteria provided by NCES.

Some special edits were used for particular types of aid. Pell award amounts were taken from the Pell program files, and divided by two if they exceeded the annual maximum of $2,300. However, if the institutional records showed a higher Pell award than the Pell program files, the discrepancy was interpreted as an initial Pell grant that had later been denied--(i.e., the institution is liable for the difference), and the excess funds were considered as institutional need-based aid. (To such students, however, it still appears that they received a Pell grant.) The maximum Stafford award amount depended on student level; also, student survey data were examined to determine whether a student might have changed status (e.g., changed undergraduate level) at some point in the NPSAS year. Perkins award amounts were compared against a list of schools authorized for Perkins loans; if a school was not authorized for Perkins loans and the student had not also attended another school in the NPSAS year (where the Perkins loan might have been legitimate), the Perkins award amount was reclassified as other federal aid. Similarly, campus-based aid at schools not in the campus-based program was reclassified as other federal aid if students had not attended more than one school during the NPSAS year. PLUS awards were examined to see if the surveyed parent said no PLUS loan was received or if the student was independent, and were reset to zero if either condition were true.

8.3.5 Statistical Imputation

NCES specified that 17 selected analysis variables should have no more than five percent missing cases. After using information from all appropriate secondary sources, there remained eight variables which required some statistical imputation. Statistical imputation generally takes two forms -- regression-based or hot deck. Table 8.3 summarizes the variables and methods used for imputation, and additional tables in Appendix C detail the imputation results for each imputed variable.

Table 8.3.--Summary statistics on imputations for the NPSAS:90 derived variable file

Variables	Total number	Data from: Record abstract	Student survey	Other	Number missing	Percent missing	Number imputed	Method of imputation
I Student characteristics								
AGE	61,120	11,676	48,056	31	1,357	2.2	0	None
MARITAL	61,120	9,764	49,703	--	1,653	2.7	0	None
HSDEG	61,120	9,599	47,535	--	3,986	6.5	3,986	Hot Deck
GENDER	61,120	11,800	47,618	--	1,702	2.8	0	None
LOCALRES	61,120	6,613	50,691	--	3,816	6.2	3,816	Hot Deck
CTZNSHP	61,120	12,501	47,150	245	1,216	2.0	0	None
RACE	61,120	7,995	45,801	2,967	4,357	7.1	4,357	Hot Deck
III Enrollment variables								
PROGTYP	61,120	60,790	306	24	0	0.0	0	None
UGRDLVL	46,788	40,435	3,637	--	2,716	5.8	2,716	Hot Deck
GPA	61,120	48,195	--	4,897	8,028	13.1	0	None*
ATTEND	61,120	57,944	153	2	3,021	4.9	0	None
IV Costs								
TUITCOST	61,120	--	2,615	58,205	300	0.5	0	None
VII Aid eligibility variables								
DEPEND	61,120	40,362	1,602	18,986	170	0.4	0	None
EFC1	61,120	29,270	--	--	32,850	52.7	0	None
SAI	13,253	702	--	12,251	300	2.3	300	Regression
DEPINC	26,232	9,523	4,242	5,554	6,913	29.2	6,913	Regression/Hot Deck
INDEPINC	34,718	10,831	7,150	921	15,816	45.6	15,816	Regression/Hot Deck

* Variables with more than 5 percent missing were imputed. EFC1 was only available for students with financial aid application data in institution records. A composite variable was created based on either EFC1, if available; or EFC2 using the CM formula. For this table, DEPINC and INDEPINC were considered "missing" if the respondent did not provide a point estimate for 1988 adjusted gross income. See Appendix C and the electronic codebook for additional information on the data imputations.

CHAPTER 9. WEIGHTS AND VARIANCE ESTIMATION

This chapter describes issues of weighting, variance estimation, and coverage of the NPSAS:90 samples. Four sets of weight variables were developed for NPSAS:90--the institution-level weights, the student abstract weights, the student-level weights, and the unduplicated student weights. In addition, a set of 35 replicate weights was developed for variance estimation purposes. The adequacy of the coverage of institutions and students within institutions were also evaluated as discussed in this chapter.

9.1 Weighting

NPSAS:90 estimation weights were produced for the analysis of data from the record abstract data and for the merged record abstract and student questionnaire data. The parent data are a supplemental source of information and not used to produce estimates and, for this reason, no estimation weights have been attached to this sample. These weights can be used to make national estimates of the number of students who were enrolled in postsecondary education in the school year 1989-90.

The methods used for producing the weights and for estimating the sampling variability of the estimates are described in the following sections. The weighting scheme for the abstract records is discussed first. The abstract weights are then used in the development of the student questionnaire weights. The replication method used to produce estimates of variance for NPSAS:90 are then discussed.

9.1.1 Institution-level Weights

The first stage of forming the estimation weights for any of the files involves using a weight appropriate for the sampling of institutions. Institutions were sampled with probabilities proportional to a measure of size assigned to the institution within the institutional sampling strata defined in Chapter 2. This probability also included the component associated with the sampling of PSU's, if the institution was not a self-representing institution sampled from the entire frame.

The inverse of this probability is the base institution weight. For notational convenience, we will refer to this weight as IW_{hjk} , where h indicates the PSU, j is the institutional stratum, and k is the particular institution. This general notation can be used to include all sampled institutions, including those selected from the Pell file and the supplemental sampling from the Stafford Loan file.

The institution weight is the inverse of the probability of selecting the institution for the sample, however, the entire sample of institutions was only used for the October data collection. Subsamples of institutions were selected for the other three time periods. To reflect the sampling across time periods, a subscript, t, is added to the institution weight for the time period.

Thus, the base institution weight is IW_{hjkt}. For t=October, the institution weight is the same as IW_{hjk}, since all institutions were included in the October sampling. For the other time periods, the base institution weight was simply multiplied by the number of institutions in the stratum divided by the number of institutions sampled for the time period.

Because the estimates of students are affected by the sampling variability associated with sampling institutions, we used a ratio adjustment procedure to reduce the variance in the estimates from this sampling stage. This adjustment was done by forming the ratio of the number of students by level (undergraduate, graduate, and first-professional) in the IPEDS Institution Characteristics (IC) and Fall Enrollment (EF) files to the estimated number of students using only the sampled and participating institutions. Thus, the enrollment data from IPEDS was used to reduce the variance from sampling institutions.

This ratio adjustment not only reduced the variance from sampling institutions, but it also was a nonresponse adjustment. The ratio can be thought of as having two components. The first component is the ratio of the number of students in the universe file (IPEDS) divided by the number of students estimated from the sampled institutions. The second component is nonresponse adjustment; it is the number of students estimated from the sampled institutions divided by the number of students estimated from the participating institutions.

For some sectors of the institutional frame, the match between the IPEDS files and the NPSAS sampled institutions was less than perfect. For example, some of the institutions sampled from the Pell and Stafford Loan files did not match the IPEDS file. For all nonmatching institutions, a separate adjustment for institution level nonresponse was computed. The adjustment used was the weighted number of eligible institutions sampled divided by the weighted number of institutions participating.

For ease of presentation we will refer to the adjusted institution weight as IW. This weight incorporates all of the adjustments done for both the matched and nonmatched institutions. It should be noted that some institutions that were originally sampled for the June 1990 data collection time period were dropped due to cost considerations. Because no students were sampled in June 1990 for some types of institutions, this elimination produces a potential bias. For other types of institutions

which were just subsampled at a lower rate than planned, the impact was to increase the variance of the estimates.

9.1.2 Student Abstract Weighting

Within each sampled, participating institution, enrollment lists were obtained and samples of students selected by student level. Because samples were selected from more than one enrollment list at a subsample of institutions, we begin by describing the basic student abstract weighting procedure and then elaborate on how this was used for multiple enrollment lists.

For each institution list, a computer file was created that contained the number of students listed and the number of students sampled, by student level (undergraduate, graduate, and first-professional). Within each of these levels (student strata for sampling purposes), a systematic sample of students was selected using a sampling interval computed based upon the previous stages of sampling.

Therefore, the sampling weight for this stage of sampling was just the number of students listed divided by the number of students sampled per stratum. The abstract weight taking account of both the institution and student sampling is AW_{hikl}, where AW is IW times the ratio of the number of listed students to sampled students in student-level stratum l.

The weighting for each time period used basically the same procedure. For October, the abstract base weights were computed exactly as described above. For the other time periods, the base weights were computed using the same procedures, but only those students who were not listed on previous enrollment lists were retained. The probability of selection was still the number of students listed (for the specific time period) divided by the number of students sampled (not the retained sample size).

Student abstracts were not completed for all of the sampled students and this nonresponse was handled by adjusting the weights for the students with completed abstracts. The nonresponse adjustment was the ratio of the number of eligible students divided by the number of students with completed abstracts. This ratio was computed by institution, student level, and time period. For most of the institutions, student level, and time periods, the ratio was equal to unity, i.e., all responded and there was no need for adjustment. If the ratio was not equal to unity and the number of eligible students was less than 30, then institutions were combined to create large enough subgroups to insure stable nonresponse adjustments.

At this stage, the student abstract weights had been subject to several different adjustments. The overall impact on the weights arising from these adjustments varied depending upon the institution, the student level, and the time period for data collection. The distribution of the weights was examined to determine if the combined effect on any of the weights would create a few cases with very large weights.

The distribution of the weights was examined within institution stratum, student level, and time period. Weights that were significantly greater than the mean weight for the group were reviewed for their impact on the variance of the estimates. Weights that were significantly greater than three times the mean weight for the group were trimmed. The trimming was done by reducing the weights of the cases with the largest weights to three times the mean weight and redistributing the difference so that the total sum of the weights for the group was unchanged by the trimming. Only a few strata and a few cases within each stratum were affected by the trimming.

After completing this phase of the weighting, examination of the results suggested that the enrollment lists of students from the time periods other than October was smaller than expected. This shortfall, which did not affect the sample size because we adjusted sampling rates, was particularly noticeable for the private, for-profit and not-for-profit, less-than-2-year institutions.

After examining, IPEDS estimates, Pell grant disbursements, and other information, the difference between the actual and expected number of students from these time periods was assumed to be related to undercoverage of enrolled students in these types of institutions for the nonfall data collection efforts. In other words, it was assumed that some of the enrollment lists did not contain all the enrolled students. Based on this reasoning, a coverage adjustment was made for the nonfall students in these types of institutions.

The coverage adjustment factor was based on the institutional stratum and it was only applied to students sampled from the nonfall time period. For 2-year private for-profit, and not-for-profit institutions, the factor was 3.0. For the less-than-2-year for-profit institutions the factor was 1.8 and for the less-than-2-year, not-for-profit institutions the factor was 1.4. These factors were empirically derived by examining the estimated number of students who received Pell Grants in these types of institutions.

A final stage of post-stratified ratio adjustments were made to bring the estimates from NPSAS into align with the estimates of the numbers from the Pell Grant file. The number of Pell Grant recipients by level and control of institution and amount of award (Pell grant award amount) was obtained by NCES.

The estimates corresponding to the Pell Grant file counts of recipients were made based on the adjusted abstract weights. The ratio of the Pell Grant totals to the NPSAS estimates was formed for each cell. These ratios were then applied to each Pell grant recipient on the NPSAS file to bring the estimates from NPSAS into conformance with those from the Pell Grant totals. These ratios are shown in table 9.1.1.

If the ratio of the totals to the estimates exceeded 2 or was less than .5, then the ratios were set at this maximum or minimum value. These restrictions were used to limit the size of the adjustment. These adjustments also caused the final estimates to differ slightly from the reported numbers in the Pell Grant file. This adjustment was made to all Pell grant recipients, regardless of the time of enrollment or their student aid status. The idea was to correct the entire file for design imperfections by this type of adjustment. The only data on which there were reliable totals for the full year was from the Pell Grant file, hence, these ratios were used to adjust all Pell grant recipients identified in the NPSAS file.

Following the post-stratification weight adjustment, the ratio of the total Stafford loan amount from NPSAS to the Stafford Loan program estimated disbursements was .96 ($8.25 billion vs. $8.56 billion). The ratio of total PLUS amount from NPSAS to the PLUS loan program estimated disbursement was .99 ($7.19 million vs. $7.29 million). The ratio of total SLS amount from NPSAS to the SLS loan program estimated disbursement was .85 ($1.47 billion vs. $1.25 billion)[6].

9.1.3 Student Questionnaire Weighting

The student record abstract file is useful for some types of analysis, but more in-depth analysis may rely on data from other sources such as the response of the students to the student questionnaire. NCES identified a file of 61,120 sampled students for which there was sufficient data for most of these types of analyses. Weights were created to support the analysis of the responses on this file of students. This file is referred to as the derived variable or *analysis file*, because these students were kept for other types of analyses.

The weighting procedures for the analysis file also began with the base weights from the institutional and student sampling within institution procedures, as outlined for the student abstract

---

[6]Note: For the Stafford Loan and PLUS programs, estimated disbursements are about 90% of loan volume commitments. For the SLS program for 1989-90, estimated disbursements are about 80% of loan commitments.

data above. These are appropriate because the institution and student sampling stages pertain to both the record abstract and the telephone interview data collection activities.

The weighting and trimming of the weights were completed using the same rules as used in the abstract weighting. The same rules were used, but, because the response patterns were different for the record abstracts and the analysis file, the application of the rules involved different values.

The nonresponse adjustment for the analysis file was very different from that used for the record abstract. There were two reasons for the difference. First, for nearly all the cases in the analysis file, some abstract record data were available to classify the students into homogeneous groups. For the record abstract weighting, no other data source was available.

Second, students were retained in the analysis file differentially depending on some characteristics of the students. For example, aided students for whom more extensive record abstract data were available, were more likely to be retained in the file than students with little record abstract data. If the nonresponse adjustment did not take the selection criteria into account, then the estimates could be subject to substantial biases.

Student nonresponse adjustment cells were set up within each institution level and control, by the following student level characteristics: student level, fall enrollment status, receipt of aid, and dependency status. For each cell, the ratio of the weighted estimate of the number of students from the entire file to the estimated number in the analysis file was computed. This ratio was then applied to the respondents in the analysis file to account for nonresponse.

The next two steps in the weighting were adjustments based on the undercoverage of the students in private 2-year and less-than-2-year institutions and based on the estimates of the number of Pell recipients. The procedures used for these two steps were identical to those used for the record abstract file. The ratio adjustments for the last step of this process are shown in table 9.2.1.

9.1.4 Unduplicated Student Weights

The weights for both the record abstract data and the analysis file can be used to produce estimates of the number of students who enrolled in an institution during the 1989-90 school year. If a student enrolled in more than one institution during the year, then that student would be counted once for each institution they attended.

This method of counting students is not unusual. For example, all of the institution-based samples or censuses, like IPEDS, use this same method. However, another method which estimates

the students uniquely, regardless of the number of institutions they attended during the year can be approximated from the NPSAS data base.

The unduplicated estimate is somewhat experimental and was done to further explore some of the issues. The weight was based on the response of the student to a question on the student questionnaire. Students who indicated that they had attended more than one institution during the 1989-90 school year were flagged. Explicit weighting that involved the chances of being selected at other institutions were difficult to develop, but a simple adjustment of the weights accomplished much of the needed modification. If the student was flagged as attending more than one institution in the year, then that student's weight was divided by two.

The weights from this procedure are not exact reflections of the probabilities of selection. Some of the shortcomings of this method are: many students may have attended more than two institutions during the year; the probability of being sampled at both institutions is not identical; and, the student might be enrolled at the institutions at different levels (undergraduate and then graduate).

Despite these potential shortcomings of the approach, the basic method should give a good approximation of the extent of the problem of students being enrolled in more than one institution during the same school year. The estimated percentage of students enrolled in more than one institution in the school year are shown in table 9.3.1 by the institution level and control and the student level.

Table 9.1.1 Ratio adjustment for abstract estimates based on pell counts, by level and control of institution and Pell grant award amount

Level and control	Pell Award Amount								
	$1-$299	$300-$599	$600-$899	$900-$1,199	$1,200-$1,499	$1,500-$1,799	$1,800-$2,199	$2,100-$2,299	$2,300
Public, less-than 2-year	2.02	1.21	0.76	0.71	0.45	1.15	1.75	0.72	0.93
Public, 2-year	1.51	1.30	0.94	1.04	1.08	0.97	0.78	0.85	0.59
Public, 4-year	0.95	0.97	0.96	0.94	0.99	0.94	0.84	1.03	0.85
Nonprofit, less-than 2-year	2.28	2.07	1.20	1.53	1.24	0.50	0.37	5.71	1.24
Nonprofit, 2-year	1.10	1.03	0.74	0.99	0.93	1.51	1.07	0.74	1.05
Nonprofit, 4-year	1.07	1.02	1.13	1.00	0.97	0.99	0.96	1.06	1.00
For-profit, less-than-2-year	1.54	1.52	1.41	1.30	1.04	1.01	1.52	1.18	0.86
For-profit, 2-year or more	0.63	1.37	0.92	1.34	0.82	0.82	0.86	0.64	0.79

Note: The minimum adjustment was 0.50, the maximum adjustment was 2.0.

Table 9.2.1. Ratio adjustment for student analysis file based on Pell grant counts, by level and control of institution and Pell grant award amount

Level and control	Pell Award Amount								
	$1-$299	$300-$599	$600-$899	$900-$1,199	$1,200-$1,499	$1,500-$1,799	$1,800-$2,199	$2,100-$2,299	$2,300
Public, less-than-2-year	0.97	1.04	0.67	0.71	0.42	1.10	1.75	0.74	0.91
Public, 2-year	1.33	1.24	0.86	1.00	1.01	0.93	0.73	0.84	0.49
Public, 4-year	0.93	0.96	0.95	0.94	0.98	0.94	0.84	1.03	0.85
Nonprofit, less-than-2-year	2.28	2.07	1.31	1.53	1.26	0.50	0.36	6.40	1.28
Nonprofit, 2-year	0.99	1.10	0.75	1.03	0.93	1.44	1.02	0.72	1.07
Nonprofit, 4-year	1.07	1.03	1.12	0.98	0.96	0.99	0.96	1.06	1.00
For-profit, less-than-2-year	1.58	1.44	1.35	1.32	1.00	1.05	1.48	1.21	0.89
For-profit, 2-year or more	0.59	1.48	0.94	1.30	0.67	0.86	0.88	0.67	0.82

Note: The minimum adjustment was 0.50, the maximum adjustment was 2.0.

Table 9.3.1.--Estimated percentage of students enrolled in more than one postsecondary institution during 1989-90

Level and control	Undergraduate	Graduate	First-Professional
Public, less-than-2-year	4.23%		
Public, 2-year	6.03%		
Public, other 4-year	5.05%	5.79%	
Public, PhD-granting	4.94%	3.68%	1.20%
Not-for-profit, less-than-2-year	1.65%		
Not-for-profit, 2-year	6.39%		
Not-for-profit, other 4-year	5.13%	5.08%	
Not-for-profit, PhD-granting	4.53%	2.61%	1.33%
For-profit, less-than-2 year	1.23%		
For-profit, 2-year or more	1.97%	0.00%	1.66%
Total	5.11%	4.00%	1.29%

9.2 Variance Estimation

The NPSAS sample design is very complex, involving several stages, clustering, and stratification of the units. In these types of complex sample designs, exact methods for computing variances are difficult to formulate theoretically and even more difficult to practically implement. For these reasons, other methods of estimating the variability of the estimates from NPSAS have been developed.

A replication method of variance estimation known as repeated jackknife replication has been used for NPSAS:90. In essence, the sample has been partitioned into 35 replicates and the estimation procedures for the full sample were repeated for each of the replicates. The procedure used to divide the sample into the 35 replicates and the method used to compute the sampling errors from the replicate estimates defines the replication method.

In the NPSAS:90, the same replication procedure was used as in the 1987 study. Each of the 34 non-self-representing strata were assigned to a variance estimation stratum and Puerto Rico, the new self-representing PSU, was assigned to the 35th variance estimation stratum. The other self-representing PSU's were allocated to the same variance estimation stratum used in 1987.

Within each of the 35 variance estimation stratum, the sampled units (PSU's, institutions, or students depending on which was the primary sampling unit in the stratum) were divided into pairs of roughly equal size. Replicates were then created by randomly dropping one unit from a variance estimation stratum and doubling the weight for the other unit in the stratum. Since there were 35 variance estimation strata, 35 replicates were formed.

For each replicate, the same estimation procedures used for the full sample were applied. This included the ratio adjustment at the institutional level, the coverage adjustments, and all levels of poststratification. Because of the complexity of the sample design and estimation procedures, the number of replicate ratio adjustments and weights that were computed in order to mimic the full sample procedures was quite large. In some instances, the final stages of poststratification to the number of Pell recipients by the Pell amounts were based on a relatively small number of records. Some collapsing to avoid null cells was necessary in these cases.

Replicate weights were created for the abstract file and the analysis file. These weights, used in conjunction with the full sample weights, were used to compute estimates of the sampling errors of the estimates. Then, the JK2 option of WESVAR, the SAS procedure developed by Westat, was used to compute sampling errors.

Another method of computing sampling errors involves approximating the estimator by the linear terms of a Taylor series expansion. This linear approximation is then used in an ultimate cluster variance estimation formula to estimate the sampling variance of the estimates. This procedure is typically referred to as the Taylor series method.

NCES has developed a computer program, called CTAB (current version is CTAB9), to estimate sampling errors using the Taylor series method. To facilitate the estimation of sampling errors for this program, we created two variables needed for its implemention. These variables were called STRATUM and PSU, and were defined by sorting and pairing the institutions by their final classification of level and control.

9.2.1 Estimates of Sampling Error

Sampling errors were estimated using both WESVAR and C-TAB. In general, the estimates of sampling errors produced by WESVAR are smaller than those computed using C-TAB. The lower sampling errors from WESVAR may be attributed to the fact that replicates used in WESVAR account for all the stages of estimation. In particular, the ratio adjustments and poststratification adjustments were included for each replicate. The C-TAB procedure does not account for these nonlinear estimation procedures. Because the impact of these adjustment procedures is to reduce the sampling error, it is reassuring that the WESVAR estimates of sampling error are smaller than those from C-TAB. Furthermore, it is typically true that the impact of the adjustments is less for sampling errors of rare characteristics and for continuous variables than for other estimates.

Table 9.4.1 gives the estimates by race, Hispanic origin, and stratum (level and control of institution) by student level. The first column is the estimated percent of all students, then the columns give standard errors of the estimates computed from C-TAB and WESVAR, and finally the ratio between C-TAB and WESVAR standard error. The estimates and standard errors are given for undergraduate, graduate, and first-professional students.

Table 9.5.1 summarizes the estimated design effects for categorical data. The table shows the unweighted number of statistics from which the summaries were computed (N) and the 25th, 50th (median), and 75th percentile of the estimated design effects. The percentiles were chosen to give a measure of the average design effect and the variability in the design effects. Means and standard deviations were examined, but they tended to distort the picture because of the presence of outliers.

The median design effect is 5.5, and the design effect is relatively consistent even when the estimates are only computed within subdomains (such as those defined by race). The first subdomain in the table is formed by dividing the estimate into three groups depending upon the size of the

estimate. If the estimate is less than 20 percent it falls into the first category, estimates between 20 and 60 percent fall into the second category, and estimates bigger than 60 percent fall into the third category. The other subdomains presented in the table have the obvious definitions.

The design effects are relatively large irrespective of the subdomain examined. Since NPSAS was specifically designed to provide reliable estimates for some rare characteristics of the population of postsecondary students, the large design effect is not surprising. The design objective was to insure that the sampling errors for certain statistics were small enough to support the required analysis, even if this required samples that were relatively inefficient for estimates of characteristics that were not rare. The standard errors of specific statistics are small, suggesting that the NPSAS design achieved this goal.

Table 9.4.1. Estimates and standard errors for categorical data in NPSAS

Categories	Undergraduate				Graduate				First-Professional			
	Estimated % of All Students	C-TAB SE	WESVAR SE	W/C Ratio	Estimated % of All Students	C-TAB SE	WESVAR SE	W/C Ratio	Estimated % of All Students	C-TAB SE	WESVAR SE	W/C Ratio
Race												
American Indian	0.68	0.07	0.08	1.18	0.04	0.01	0.00	0.25	0.01**	0.00	0.00	--
Asian	4.10	0.28	0.20	0.72	0.92	0.06	0.04	0.69	0.12	0.01	0.01	0.85
Black	8.94	0.60	0.66	1.10	0.57	0.05	0.04	0.76	0.07	0.01	0.01	0.53
Hispanic	7.37	0.50	0.35	0.69	0.47	0.04	0.03	0.73	0.08	0.02	0.01	0.74
White	66.44	0.90	0.80	0.89	8.85	0.46	0.15	0.32	1.34	0.10	0.02	0.20
Hispanic												
Yes	7.01	0.49	0.32	0.66	0.44	0.04	0.03	0.74	0.08	0.02	0.01	0.71
No	74.84	0.72	0.47	0.65	9.54	0.49	0.16	0.32	1.42	0.10	0.02	0.23
Stratum												
Public, < 2-year	1.24	0.29	0.29	0.98								
Public, 2-year	36.69	1.56	0.41	0.26								
Public, other 4-year	12.35	0.90	0.48	0.53	2.24	0.23	0.09	0.39				
Public, Ph.D.	15.95	1.29	0.43	0.33	4.88	0.51	0.14	0.27	0.63	0.06	0.01	0.10
Private, < 2-year	0.39	0.10	0.09	0.89								
Private, 2-year	1.06	0.12	0.05	0.38								
Private, other 4-year	7.86	0.46	0.26	0.56	1.13	0.14	0.15	1.04				
Private, Ph.D.	4.51	0.35	0.22	0.62	2.60	0.18	0.11	0.60	0.96	0.09	0.01	0.13
Proprietary, < 2-year	4.95	0.48	0.15	0.32								
Proprietary, 2-year	2.54	0.33	0.09	0.28					0.04	0.09	0.00	0.02

* Note, this summary was updated to reflect corrections made to the file.
** Because the percentage of American Indians is so small, the ratio between WESVAR and C-TAB is not given.

Table 9.5.1. Design effects for categorical data in NPSAS

	N	Q1 (25%)	Median	Q3 (75%)
All	568	3.0	5.5	9.9
Estimate Groups				
Low - < 20 percent	486	2.8	5.1	9.8
20 - < 60 percent	66	5.5	7.6	10.9
60+ percent	16	2.8	5.0	10.3
Sample Size				
< 2,000	71	1.3	3.1	4.9
2,000 - < 10,000	121	2.0	4.2	10.3
10,000+	376	4.1	6.3	10.3
Race				
American Indian	72	2.6	4.4	5.8
Asian	72	2.5	3.9	6.1
Black	72	4.7	10.3	15.6
Other	72	3.4	6.2	10.4
White	72	4.3	7.6	12.2
Hispanic				
Yes	73	3.5	6.2	9.9
Stratum				
Public, < 2-year	63	2.9	7.1	16.0
Public, 2-year	63	2.9	5.8	10.3
Public, other 4-year	63	3.2	6.4	10.3
Public, Ph.D. or first prof.	63	3.2	5.9	8.0
Private, non-profit < 2-year	63	2.3	5.0	10.3
Private, non-profit 2-year	63	1.3	2.5	5.9
Private, non-profit other 4-year	63	3.1	5.8	10.1
Private, non-profit Ph.D. or first prof.	63	2.7	4.3	5.9
Private, for-profit < 2-year	63	2.6	3.7	8.1
Private, for-profit 2-year or greater	63	2.5	3.6	7.9

9.3 Coverage Issues

An important issue for every survey is the coverage of the target population, i.e., the extent to which portions of the population for which inferences are desired have been included in the sampling frame. If units (students or institutions) are excluded from the sampling frame, then they have no chance of being in the sample. Inferences intended for the entire population are actually valid only

for that portion of the population which is included in the sampling frame, unless adjustments are made to account for the undercoverage.

Estimated totals are less than the expected totals when some portions of the population are excluded from the sampling frame. For example, if postsecondary institutions in urban areas were undercovered in NPSAS, then estimates of the number of students from urban areas would be less than expected. Means and proportions may also differ from the expected values, depending on the extent of undercoverage and the differences in the characteristics of the covered and omitted populations. For example, if part-time students were undercovered in NPSAS and the cost of attendance for these students were substantially lower than for other students, then the estimated mean cost of attendance from NPSAS would be greater than expected.

The coverage of postsecondary students in NPSAS:90 can be divided into two separate issues: the coverage of postsecondary institutions and the coverage of postsecondary students within those institutions. In both cases, undercoverage can arise for a number of reasons. For example, undercoverage might exist due to the following:

- the frame from which the sample of institutions was selected may exclude some types of institutions,
- the institutions which have recently been established may be excluded,
- the list of students provided by the institutions may not include off-campus students, and
- the list of nonfall students may be matched incorrectly against the list of previously enrolled students.

9.3.1 Institutional Coverage

The coverage of institutions for NPSAS:90 is analyzed below using data from IPEDS and the Pell Grant Institution files. We conclude that there is no evidence of a large bias associated with undercoverage of institutions, but that the issue of eligibility of institutions for NPSAS needs further study. The conclusions are tentative due to other errors that confound the analysis, especially the evaluation with respect to the Pell file.

The comparisons below examine the coverage of postsecondary students as the result of the completeness of the sampling frame of institutions. In other words, the number of students is used as a measure of the completeness of coverage. If a few large institutions were excluded from the sampling frame of institutions, the undercoverage of students might be worse than the exclusion of a

BEST COPY AVAILABLE

larger number of small institutions. Because of the analytic importance of the level and control of the institution, the coverage is examined separately by these characteristics.

IPEDS Comparisons

One of the sources that can be used to evaluate the coverage of institutions in NPSAS:90 is the IPEDS IC file for 1989-90. Two important factors must be considered when evaluating institutional coverage using the IPEDS file. First, since the 1987-88 IPEDS IC was used as the primary source for the sampling frame for NPSAS:90, our assessment of the coverage with respect to the 1989-90 file might overestimate the coverage of NPSAS:90. However, the IC file underwent major changes between the 1987-88 file and the 1989-90 file and this suggests that the assessment is worthwhile. The 1987-88 file consisted of about 12,100 institutions while the 1989-90 file had under 10,900, including about 10,300 institutions that were on both files. Less than 500 new institutions were added and 1,800 were deleted between the those years. The deletions were concentrated in the less-than-2-year institutions.

Second, the IPEDS IC file is not absolutely complete for either year and should not be considered to contain the "true" values. The comparisons against the IC file are informative, but not definitive. The suspected lack of completeness of institutions in IPEDS prompted the use of the Pell and Stafford loan files for the construction of the sampling frame originally. Of course, no completely accurate count of postsecondary institutions exists, and IPEDS was used because it was considered to be the most complete and accurate national file.

The following types of institutions were deleted from the IC file before the analysis:

- institutions outside the 50 states, D.C., or Puerto Rico,
- institutions that were only systems offices,
- federal institutions (mainly service academies), except for Indian schools,
- a few institutions that had only correspondence students, and
- institutions that had no undergraduate students, (those with only graduate or first-professionals enrolled).

In all, about 200 institutions were excluded for one or more of the reasons above. Only data on fall, undergraduate students are used in this analysis.

Comparing the undergraduate enrollments from the IC file to the estimated number based on the sampled NPSAS institutions is one measure of institutional coverage. The ratio of the estimated enrollment from the sampled institutions to the universe count is 1.01 (1.07 when the unmatched institutions are also included in the total), indicating that overall there is no apparent undercoverage. However, the ratios for the following types of institutions are under 90 percent: public, less-than-2-year (.75); private, not-for-profit, less-than-2-year (.74); private, not-for-profit, 2-year (.80); private, for-profit, less-than-2-year (.85); and private, for-profit, 2-year or more (.84). The sampling errors for the first three categories are probably substantial because the sample size for each category is relatively small.

These lower ratios can be attributed to three factors: new institutions on the IC file that were not included in 1987-88 IPEDS, the inability to match institutions because of name changes and other assorted matching problems, and sampling errors in the estimates. The first of the three factors is associated with institutional undercoverage. In general, the undercoverage of institutions, even in the for-profit sector, does not appear to be severe from this perspective.

The undercoverage can be viewed differently by restricting the ratios from the NPSAS sample to those institutions which were designated as being eligible for NPSAS. The three main reasons for ineligibility in NPSAS:90 were: institutions only had programs lasting less than 3 months or 300 clock hours, institutions only offered correspondence programs, and institutions only offered GED and/or remedial courses.

Comparing the estimates from the eligible institutions to the sampled institutions shows the importance of the eligibility rules, especially for some of the categories of institutions. Overall, the ratio of the estimated total enrollment dropped 6 percent. The categories in which the ratio was less than 98 percent were: private, not-for-profit, less-than-2-year (.58); private, not-for-profit, 2-year (.89); private, for-profit, less-than-2- year (.63); private for-profit, 2-year or more (.75), and the unmatched category (.51).

Thus, the eligibility of institutions had a more significant impact on the estimated enrollment than the coverage problems with respect to IPEDS, especially for the private 2-year and less-than-2-year sectors. Almost all the institutions from the other sectors were eligible. The unmatched institutions also included a large proportion of ineligible institutions.

Institutional coverage with respect to IPEDS can also be examined by basing the estimates on the institutions that participated in NPSAS:90. The weights used to produce these estimates were the base weights multiplied by a nonresponse ratio adjustment factor. The estimates of undergraduates for most categories of institutions were larger than the estimates based on the eligible institutions. The

categories of institutions where the estimates decreased (public, 2-year; public, doctorate-granting; private, not-for-profit, 4-year; and private, not-for-profit, doctorate-granting) are those which contain the majority of students.

Another way of viewing the impact of the sampling and weighting of the institutions for NPSAS:90 is by computing the ratio of the estimates of the number of undergraduates from the participating institutions to the counts from the universe (see table 9.6.1).

Table 9.6.1.--Ratios of the estimated weighted number of undergraduates in NPSAS eligible institutions to the number of undergraduates as reported in IPEDS

IPEDS Institution Level and Control	Ratio
Public, less-than-2-year	0.98
Public, 2-year	0.98
Public, other 4-year	0.99
Public, doctorate-granting	1.00
Private, not-for-profit, less-than-2-year	0.52
Private, not-for-profit, 2-year	0.88
Private, not-for-profit, other 4-year	0.98
Private, not-for-profit doctorate-granting	0.99
Private, for-profit, less-than-2-year	0.59
Private, for-profit, 2-year or more	0.74

These ratios highlight the fact that, even after adjusting the weights, the private (both for-profit and not-for-profit), less-than-2-year and 2-year institutions are below the IPEDS totals. Because the sample size and the population size of the not-for-profit sector is small, the major problem is in the for-profit sector. In this sector, the main reason for the decrease in the estimates was the eligibility of the institutions, not institutional coverage.

Pell Grant File Comparisons

The other data source that can be used most easily to examine the institutional coverage in NPSAS:90 is the file of institutions participating in the Pell Grant program. The 1989-90 Pell Institution file (a preliminary file created January 28, 1991 and provided to us by the Pell Grant Branch) was used for this evaluation. The file contains the number of recipients and the expenditures for each institution participating in the Pell Grant program.

This evaluation was conducted using the same procedures as used with the IPEDS comparison. The sampled NPSAS institutions were matched to the Pell Grant institution file and the characteristics from the Pell file were used to produce the estimates. Once again, these estimates do not reflect data collected in NPSAS:90; nor do they match the students sampled for NPSAS:90 with a file of Pell grant *recipients*. The Pell Grant institution file only contains institution-level aggregates for each institution. The estimates will not correspond to any of the estimates derived from NPSAS:90 for this reason. The tables should only be used to evaluate the coverage of institutions.

The evaluation of institutional coverage is complicated somewhat because of the errors associated with matching institutions between the two files. The IPEDS matching was much simpler because the institutions had a common identifier on them. Matching was further complicated by the fact that the two data systems have different administrative rules that confound the matching process. For example, a single institution may report for institutions located in more than one location in one file, but all the locations may report separately in the other file.

The overall impact of the matching errors suggests that the institutional coverage for NPSAS:90 is worse than it actually is. This conclusion is based on the assumption that errors from the inability to match institutions that are the same exceeds the error of matching institutions that are not the same. The assumption seems reasonable and is tentatively supported by the fact that students with Pell grants were found in 46 institutions that remained unmatched despite several levels of review.

The Pell Grant Institution file contained nearly 9,000 institutions, but only 6,873 of them have Pell grant recipients. During the matching process, a NPSAS sampled institution could be matched to an institution on the Pell file irrespective of whether Pell grants were awarded at the institution in 1989-90. A total of 181 sampled institutions matched to those on the file without Pell grant recipients and 116 of those sampled institutions participated in NPSAS:90.

The ratios of the sampled institutions (both recipients and expenditures) to the universe are shown in table 9.7.1. All of the ratios exceed .80, except the public, less-than-2- year; the private, not-for-profit, less-than-2-year; and the private, not-for-profit, 2-year. These three categories account for only 3 percent of the number of recipients, indicating that both sampling error and matching errors could account for a substantial portion of the apparent problem.

Table 9.7.1--Ratio of NPSAS estimates to Pell grant program totals, by institution level and control

Pell Institution Level and Control	Ratio of Sample to Pell Universe	
	Recipients	Expenditure
Public, less-than-2-year	0.75	0.69
Public, 2-year	0.91	0.89
Public, other 4-year	1.01	0.95
Public, Ph.D. granting	0.89	0.84
Private, not-for-profit, less-than-2-year	0.43	0.40
Private, not-for-profit, 2-year	0.66	0.68
Private, not-for-profit, other 4-year	1.02	0.92
Private, not-for-profit Ph.D. granting	0.92	0.89
Private, for-profit, less-than-2-year	0.93	1.00
Private, for-profit, 2-year or more	0.87	0.83

The ratios suggest that the problem is not especially concentrated in one or a few categories, but rather evenly distributed over all the categories which have a substantial number of Pell grant recipients. The ratios for the total number of recipients and expenditures are 0.91 and 0.88, respectively.

Given the fact that the ratios are relatively consistent across categories of level and control, two hypotheses seem feasible. One possibility is that NPSAS:90 does not fully cover eligible institutions in the Pell file, but this undercoverage is not very dependent on the level and control of the institution. A second hypothesis is that the consistency of the ratios indicate that matching error is greater than undercoverage, but the matching error is not much more pronounced in one category of institution than another.

Without eliminating the matching error, it is not possible to distinguish among competing hypotheses to explain the ratios. Having worked with the files of institutions, we tend to believe that matching error is a significant problem and may obscure the ability to see coverage problems.

The estimates for the eligible sampled NPSAS institutions were also examined. Since an institution with Pell grant recipients should be eligible for NPSAS in all but a few circumstances, the close correspondence between the estimates was expected. The only category that posed a new problem was the private for-profit, less-than-2-year institutions. The number of recipients decreased by about 6 percent due to eligibility requirements.

Table 9.8.1 shows the ratios of the estimates from the participating institutions to the universe, and gives a more complete view of the institutional coverage.

Table 9.8.1.--Ratios of the Pell grant estimates from NPSAS participating institutions to the Pell grant universe

Pell Institution Level and Control	Ratio of Sample to Pell Universe	
	Recipients	Expenditure
Public, less-than-2-year	0.89	0.81
Public, 2-year	0.91	0.89
Public, 4-year	1.12	1.05
Public, 5-year or more	0.85	0.80
Private, not-for-profit, less-than-2-year	0.55	0.59
Private, not-for-profit, 2-year	0.80	0.82
Private, not-for-profit, 4-year	0.80	0.70
Private, not-for-profit, 5-year or more	0.88	0.86
Private, for-profit, less-than-2-year	1.01	1.02
Private, for-profit, 2-year or more	0.99	0.93

These ratios again reveal that the analysis does not suggest undercoverage in specific categories of institutions, but a more consistent trend across categories. The nonresponse adjustment, which was based on enrollment counts rather than Pell recipients or awards, has made considerable improvements in the estimates. The private for-profit, less-than-2-year category, which is the category thought most likely to suffer from institutional undercoverage, has ratios exceeding 1.00. This supports the notion that matching error and sampling error--rather than undercoverage--are more likely to be the reasons for the low ratios noted earlier.

9.3.2 Within Institution Coverage

The other potential source for coverage errors is the enrollment list provided by the institution for sampling purposes. As noted earlier, some segments of the student population could have been inadvertently omitted from the list. Such omissions may be associated with certain types of students, such as students who take courses off-campus, those enrolled in continuing education programs, part-time students, etc.

The coverage of the student lists is compounded by the need to cover students across the entire school year. Thus, the list of students who were enrolled in October 1989 could have been very complete and accurate, but if the lists provided for the other time periods were not complete then undercoverage would still be a problem. The exclusion of certain types of schools from the June sampling means that those students who were not enrolled in the school earlier in the year were not covered by NPSAS:90. The decision to exclude this sampling was based partially on the assumption that the udercoverage from this source would be minimal.

The evaluation of the undercoverage of the enrollment lists used for sampling students for NPSAS:90 is very difficult to accomplish. Conceptually, the undercoverage for a specific institution could be evaluated by taking the enrollment lists supplied by the institution, extracting a unique (each student listed only once) enrollment list for the year, and comparing this list to a known complete count for the school year. Unfortunately, this is not possible; in part, because an accurate unduplicated headcount of students enrolled at any time during a given period (e.g., enrolled at anytime between July 1, 1989 and June 30, 1990) for each postsecondary institution does not exist.

The first problem is that only 53 of the 1,130 participating institutions were sampled and provided lists for all four time periods. This reduction in the number of lists required from the schools was done by design to ease the burden on the responding institutions. The elimination of some of the sample from the June sample means that complete enrollment lists for the entire year are not available for virtually any of the 4-year institutions.

A second problem is that the enrollment lists provided by the institutions are not easily matched. Many of the lists were provided only in hard-copy format and the order of listing was different from one time period to the next. The production of a unique list from this source would be extremely error-prone and expensive.

An alternative to matching is to compute estimates of the number of students based on the sampled students. This approach is attractive, but errors made in matching the students sampled (the unduplication done during the sampling process) would be included in these estimates and the sample errors on the totals would be relatively large for many institutions.

Both of these approaches to evaluation presume that a complete, accurate, unduplicated head count of the number of students enrolled in the institution is available for comparison purposes. Even this is not, in general, the case. The only counts which purport to be complete unduplicated head counts are now available from IPEDS. However, the IPEDS procedure for obtaining these counts is new and has not yet been evaluated.

As a result of these difficulties, quantitative evidence on the within-institution coverage for NPSAS:90 is not available. Evaluation efforts using the available data raises more issues than it resolves. The potential for undercoverage in estimates due to the enrollment lists is real and needs to be addressed. In addition to examining the data collection sources of errors, efforts to evaluate other head counts for postsecondary institutions would be most useful.

CHAPTER 10. SUMMARY OF THE 1990 NPSAS FIELD TEST

A field test, in general, serves to test alternative procedures or instruments that might be used in a full-scale survey. For example, a field test can be used to test the effect of a range of monetary incentives on response rate or to validate the phrasing of a particular survey item or to judge the overall reliability of responses. Results from the 1990 NPSAS field test served as the basis for some specific decisions about the full-scale survey. (See Figure 10.1 for a summary of field test results and Figure 10.2 for the basic elements of the field test design.) The specific purposes of the 1990 NPSAS Field Test (FT) were to:

- examine the reactions of participating institutions to the overall objectives, sampling strategies, data collection procedures, and plans for testing new students in the full-scale NPSAS, based on a **pretest** of nine institutions;
- develop **field test cohorts** for new students and for seniors, to be used in related future NCES longitudinal studies;
- improve the comprehensiveness and accuracy of **financial aid data collected by the record abstractors** for graduate and first-professional students
- test the **effectiveness of computer aided telephone interviewing** when surveying postsecondary students
- improve methods of **obtaining college entrance exam test scores** or other standardized measures of ability

Figure 10.1--NPSAS:90 Field Test Results

Procedure or Design Issue Tested	Result
1. Pretest/clinical trials	1. Successfully identified problems with overall design that were confirmed through the field test.
2. Develop Cohorts for First-Time Students	2. Schools could not accurately identify new students since they do not uniformly request transfer information.
3. Record Abstraction for Graduate/ Professional Students	3. Data from individual graduate departments should be sought.
4. Computer-Assisted Telephone Interviewing	4. Very successful method of obtaining student and parent information, if good training, rigorous locating procedures, and careful administration.
5. Obtaining Standardized Test Scores	5. If test scores unavailable, schools could offer "drop-in" testing sites to accommodate a students' schedules.

Figure 10.2--NPSAS:90 Field Test Design

Survey Design Component	Description
Frame	IPEDS Postsecondary Institutions
Sample Selection	Institutions sampled from eight geographic areas A purposive sample of 75 schools (see Figure 1)
Sample Size	FT, FY 2,753 Seniors 1,036 Graduate 502 First-Prof. 210 Total 4,501
Instruments	CATI survey instruments with 10 main sections: -- school enrollment -- enrollment status and costs -- financial aid -- other sources of support -- employment -- demographics and plans -- parent characteristics -- student dependency and financial status -- longitudinal baseline data -- locating information
Data Collection	• Scripted and role-played training approach • Record abstractors and 96 telephone interviewers • Overall response rate: 73 percent
Data Processing	• Data entry
Reports	• Field Test Methodology Report

Pretest

For the pretest prior to developing materials for the field test, we selected a purposive sample of nine postsecondary institutions, based on location, type and control, size and calendar system. Two schools had also participated in the previous NPSAS, enabling these trials to examine the effect of earlier surveys on current data collection methods. (See table 10.1 for a summary of pretest school characteristics.)

Table 10.1--Schools in Pretest

Control Level	Location	Calendar System	Enrollment	Visit Date
Public, 4-year doctoral	MD	4-1-4	4,500	1/25/89
Private, 4-year doctoral	DC	Semester	6,700	1/5/89
Public, other 4-year	WV	Quarter	3,900	12/19/88
Public, other 4-year	PA	Semester	7,200	2/2/89
Public, other 4-year	VA	Semester	3,300	12/15/88
Private, other 4-year	MD	Semester	1,900	1/18/89
Public, 2-year	MD	Semester	13,000	12/14/88
Private, for-profit, 2-year	MD	Trimester	480	12/20/88
Private, for-profit, less-than-2-year	MD	Program-specific	45	2/21/89

NPSAS staff contacted sampled schools and made arrangements for site visits, to be conducted jointly by Westat and NCES. Participants were assured that no data would be collected during these trials, only discussions and reviews of data collection procedures.

Site visits took the form of discussion sessions, led by Westat staff, which explored the views of school staff on the best means to collect data in several areas. Major points of agreement on these subjects are summarized below:

Enrollment Lists and Headcounts

- Accurate headcounts should be obtained from the registrar or the director of institutional research, or the party responsible for completion of the school's IPEDS and FISAP forms.

- Most schools (6 of 9) could supply printed lists of students, if given a month's notice.

-- Traditional calendar schools typically finalize their enrollment lists by the first week in October or March. Other schools generally have continuous registration, but usually update their enrollment lists after the third class meeting.

-- It is probably impractical to require unduplicated lists from any but the most technologically-sophisticated institutions.

Current Addresses for Students and Parents

-- Generally, institutions felt confident of identifying a currently enrolled student's permanent address, but less confident of tracking a student's local off-campus address. Graduates might best be tracked through the alumni office.

-- Parents' address(es) are not regularly collected by the institution. For aided, dependent students one source will be the aid application. For nonaided students, older students, independent students, and others, the task of identifying an address for a parent, if none is available through the student, will be more difficult.

Record Abstract Data Collection

-- Generally, schools have a central financial aid office which has computer access to most financial aid data for each student. Some key exceptions to aid data availability include: **graduate assistantships** which are not defined as aid and, as a result, may only appear on records in the individual graduate departments; **veteran's benefits** which is recorded by the veterans benefits coordinator at the school; and **private scholarships** which are not reported to the school.

-- Since schools issue 95 percent of their awards by the beginning of the fall term, major data collection could occur in the fall. Full data collection would, however, require an update later in the year.

Testing or Obtaining Test Scores for New Students

-- Schools could not guarantee accurate identification of first-time, first-year (new) students because they do not ascertain previous postsecondary attendance, unless credits are transferred.

-- Aside from identification, 4-year schools could easily provide test scores for their new students but 2-year and less-than-2-year schools, which do not require these tests for admissions, would likely not have scores available.

-- For those for whom test scores are unavailable, testing on campus appears to be a somewhat intractable alternative. Schools suggested that students would not show up for tests even if monetary incentives were offered and that many students who work while in school simply have no time to take these tests. The only positive suggestion was to offer a "drop-in" test site which could accommodate students at their convenience.

Developing Field Test Cohorts

Initially, NPSAS:90 was to form the basis for two longitudinal cohorts, one for beginning students, the other for graduating seniors. Budget constraints, however, required the postponement of testing and developing a senior cohort. Thus, the field test objective in this area was narrowed to a determination of a school's ability to identify first-time, first-year students, i.e., those that had never before attended postsecondary education.

Even before carrying out the regular field test, the pretest revealed that schools could not unequivocally identify FT-FY students. Each of the institutions felt that they could identify students who were attending their institution for the first time. However, previous postsecondary attendance would only be known if the student was transferring credit. Representatives from traditional 4-year schools were confident that the majority of their first-time students were either new to postsecondary education or would report credit transfers, but allowed that many exceptions could be found. The less-than-2-year and 2-year institutions do not even ask about previous postsecondary education. Community colleges especially would have trouble with this item since many of their students are college graduates taking courses for enrichment.

Our recommendation based on the pretest results was to obtain our sample of FT-FY students from 4-year schools only. We did not test this finding in the regular field test as we believed that the trial findings would clearly be borne out.

Abstracting Financial Aid Data for Graduate and First-Professional Students

For this part of the field test, we analyzed the difficulties in collecting financial aid data for G&FP students by testing a sample of 708. Based on the 1987 NPSAS experience, it was expected that special efforts would be necessary to obtain complete aid data for these students. Toward this end, our data collectors were instructed to visit departmental and other noncentral offices at the institution to ensure complete coverage of potential sources of aid data. Data collectors reported high levels of cooperation at the schools in obtaining aid data from these noncentral offices. Therefore, it is unlikely that our data collection method needs to be modified.

However, based on data collectors' comments, the record abstract instrument could be improved in three ways:

- within the Institutional Aid section, allow abstractors to record specific names for "need-based grants," "non-need-based nonfederal scholarships" and "other-specify;"
- separate the reporting of assistantships and scholarships from the institutional aid section; and
- eliminate the possible response "not specified" because it proved not useful.

Computer-Assisted Telephone Interviewing (CATI)

How well would CATI work when administering the NPSAS survey? What improvements could be made for the full-scale survey? It was necessary to obtain answers to these questions in the field test, because NPSAS:90 was to be administered solely by phone. There were three areas for which we identified specific improvements:

- training
- locating
- administering

Training

The CATI training program used scripts of the CATI questionnaire and portable computers to teach concepts to the 96 interviewers we trained. The first script was presented by a project staff member using an interactive role-playing technique whereby trainees take the role of interviewer and the lecturer acts as respondent, emphasizing various points and providing specific instruction when necessary. As the trainees recorded their answers into the CATI screens, they were instructed to check their recordings against the correct recording on one of two overhead electro-screens. More complicated examples and instruction were then presented. Again the trainees took the role of interviewer, and a member of the project staff, using a script, played respondent. Interspersed with the scripts were exercises designed to reinforce some of the more difficult concepts in the questionnaire. After the interactive lectures, role-plays were done in pairs. Each pair of interviewers completed at least two scripted role-plays. With the first script one member of the pair played the role of interviewer while the other was respondent. The trainees changed roles with the next script. The purpose of the role-plays was to provide additional practice with the CATI questionnaire and allow the trainee to get a feel for the flow of the interview without being interrupted.

Telephone interviewers reported that role-playing exercises were the most useful part of training for actual interviewing. We recommend that this aspect of training be enhanced for the full-scale study and that exercises be developed using actual examples from the field test of both typical and unusual situations.

We also recommend that consideration be given to developing a Spanish version of the student CATI survey for administration to students in Puerto Rico.

Locating

There were significant problems locating students for the field test. Of the 4,501 sampled students eligible for the student survey, 373, or 8.3 percent, were unlocatable. An additional 42 were ineligible, 426 were refusals, and 404 were other nonresponses, leaving 3,256 completed interviews. The unlocatables could not be reached at any of the phone numbers provided by the school or through directory assistance. For 200 of these a commercial locating service was engaged to search for additional locating information. This, however, proved ineffective resulting in only 12 new completes.

We could not use DMV searching, generally the most effective in tracing, because date of birth was not generally available in the field test. This will not be the case in the full-scale study and so should help reduce the percentage unlocatables.

We therefore recommend using a combination of commercial locating and DMV search services to maximize our full-scale study response rate.

Administering

The student survey of NPSAS:90 was conducted using CATI. Through CATI, the survey instrument is programmed using specialized software and program-controlled displays are produced which guide the interviewer through the questionnaire. The data file is created as the telephone interviewer enters responses during the interview. With CATI, minimal additional editing of the data is required prior to file production, and followup calls to respondents for critical item retrieval are not necessary.

All skip patterns and range, format and consistency edits are programmed into the system. The interviewer can immediately ask the respondent to clarify or correct responses that the computer flags as questionable. For example, in Section 3 of the Student Survey an error message is displayed for the interviewer if more than a 10 percent discrepancy exists between total financial aid figures reported by the respondent (e.g., total grants) and the total calculated by the computer based upon prior respondent-reported amounts (e.g., how much for each scholarship or grant). Then the interviewer can explain the discrepancy to the respondent, who can reconcile the information immediately.

Other features of the CATI system include: (1) a data dictionary/codebook with variables ranges, formats, record layouts, and labels; (2) capability to create and process hierarchical file structures to eliminate data redundancy and conserve computer resources; (3) a scheduler system to

manage the flow and assignment of cases to interviewers by time zone, appointment, etc.; (4) automatic audit file creation to ensure that if an interview is prematurely terminated and later restarted, all data entered during the previous interview attempt can be retrieved; and (5) a screen library containing the survey instrument.

The following recommendations are based on our phone interviewers' experience with Student Survey CATI interviewing:

- matrix screens should be used to collect and verify information where possible in the attendance, cost and employment sections of the Student Survey instrument;
- when available, selected data obtained on the Record Abstract should be loaded into the CATI system to drive skip patterns by eliminating the duplicate collection of information already available on the Abstract;
- precoded categories should be used wherever possible to avoid open-ended responses or literal responses that need to be coded later; and
- for the full-scale study proxy items should be asked of respondents who cannot provide detailed 1040 tax form information.

Obtaining Ability Measures

Another purpose of the field test was to test methods of obtaining college entrance exam scores or other standardized measures of ability.
Specifically we examined two areas:

- the availability of SAT/ACT scores for FT-FY students; and
- failing this, the ability of institutions, using monetary incentives, to attract first-time students to take an Aptitude Test.

Availability of SAT/ACT Scores

The pretest also revealed that only 4-year schools could easily or at all provide entrance exam scores. The 4-year doctoral and other 4-year institutions that require and/or collect these scores would have no problem providing them to data collectors. They are usually not kept in the computer files, but could be obtained by searching hardcopy in the admissions office. The 2-year and less-than-2-year schools do not require SAT/ACT scores for admission, although if a student provides one it would probably be in the file.

Therefore, we decided to assume availability at 4-year schools and to test the effect of monetary incentive on aptitude test taking at 2-year and less-than-2-year schools only.

Administer NPSAS Aptitude Test (NAT) to First-Time Students

The goal of this portion of the field test was to administer an aptitude test (the so-called NAT) to approximately 1,300 first-year first-time students and to test response rates using monetary incentives. This test was to serve as a substitute measure of ability for those students who had not taken or did not report college entrance exam scores (e.g., SAT, ACT). Most of these students without scores are likely to be enrolled in 2-year and less-than-2-year schools that usually do not require SAT or ACT scores for admission. Consequently, we drew our sample from them.

We sampled 1,534 students in 2-year and less-than-2-year schools and offered either no incentive or a $10 or $20 incentive to each to appear for an Aptitude Test. Our overall response rate was an unexpectedly low 40 percent, yielding only 607 completed tests. Response rates by incentive level were: 33 percent for no incentive, 31 percent for $10, and 55 percent for $20. We found that a large number of appointments for test administration were made and then broken. The level of broken appointments suggests that a substantial share of students were averse to taking these tests given the incentives offered.

An analysis of response rate controlling for class time (whether taken during class or otherwise) and institutional control (public vs. private) yielded the following estimated response rates by incentive level:

Estimated Probability of Response	Sample Group
47%	All
34%	No incentive
39%	$10 incentive
68%	$20 incentive

We did not achieve the anticipated number of test takers due to a very low response rate. To maximize response, we would recommend two modifications:

- offer only a $20 incentive; and
- seek institutional support to permit students to take test during class time.

Even with these measures it is likely that we would not meet NCES' response rate standards. Therefore, we would recommend weighing the costs of obtaining higher response rates against the benefits of ability testing.

Appendix A

Sample Materials Mailed to Sampled Institutions and Students

UNITED STATES DEPARTMENT OF EDUCATION

OFFICE OF THE ASSISTANT SECRETARY
FOR EDUCATIONAL RESEARCH AND IMPROVEMENT

National Center for Education Statistics

June, 1989

Dear Chief Administrator:

The National Center for Education Statistics (NCES) of the U.S. Department of Education is continuing the National Postsecondary Student Aid Study (NPSAS), a major national study on student financial aid. The first NPSAS was conducted during the 1986-87 school year and involved more than 1,000 postsecondary institutions. In response to the continuing need for the comprehensive data provided by NPSAS, Congress has legislated that the study will be repeated every three years.

NCES will conduct the second cycle of NPSAS during the 1989-90 school year. A number of enhancements have been made to the design of this second NPSAS, including coverage of students who attend school at all different points of time throughout the year, and enhanced collection of graduate and first-professional student information. In addition, first-time students and graduating seniors will provide base year data for subsequent longitudinal studies. Further details of the 1990 NPSAS design are presented in the enclosed Study Summary.

The 1990 NPSAS will involve about 1,200 postsecondary education institutions and about 70,000 students. Your institution has been selected to participate in the study this year. As a NPSAS participant your institution will receive a copy of summary reports and tabulations. Because of the way in which the sample of institutions was selected, your participation is very important to the success of this effort.

We are writing to request that you appoint a coordinator to serve as a point of contact for the 1990 NPSAS. Staff from Westat, Inc., our contractor, will work with your coordinator to arrange for data collection in an efficient and convenient manner at your institution. Enclosed is a postcard on which to provide your coordinator's name and title.

During the study, a NPSAS representative will be available to visit your institution to assist with data collection. Your institution will be asked to provide lists of students enrolled during the 1989-90 school year from which a sample can be drawn, and to provide certain data on the sampled students. Further details of the data collection process, our assurance of confidentiality of the data, and estimates of time commitments of your institution during the study are contained in the enclosed information materials.

We appreciate your cooperation. If you have any questions about the study or our procedures, please contact either the Westat Director of Survey Operations, Ms. Diane Ward (800-937-8288) or one of the NCES Co-Project Officers, Dr. Sandra Garcia (202-357-6597) or Mr. Gerald Malitz (202-357-6364).

Sincerely,

Emerson J. Elliott
Acting Commissioner

Enclosures

WASHINGTON, D.C. 20208

1990 NATIONAL POSTSECONDARY STUDENT AID STUDY

(1990 NPSAS)

Conducted by Westat, Inc.
for the National Center for Education Statistics,
U.S. Department of Education

The professional organizations listed below endorse the
1990 National Postsecondary Student Aid Study and encourage
their members to cooperate in this important study.

American Association of Collegiate Registrars and Admissions Officers (AACRAO)

National Association of Student Financial Aid Administrators (NASFAA)

American Association of Community and Junior Colleges (AACJC)

American Council on Education (ACE)

Council of Graduate Schools (CGS)

Association of Independent Colleges and Schools (AICS)

National Association of Accredited Cosmetology Schools (NAACS)

National Association of Trade and Technical Schools (NATTS)

National Institute of Independent Colleges and Universities (NIICU)

1990 NATIONAL POSTSECONDARY STUDENT AID STUDY
INSTITUTION INFORMATION SHEET

The 1990 National Postsecondary Student Aid Study will be conducted during the 1989-90 school year. Included in the 1990 NPSAS will be record data collection at about 1,300 postsecondary institutions and surveys of approximately 70,000 students, and 26,000 parents of students. NCES' contractor, Westat, will provide field staff to carry out the data collection activities at your institution. Listed below is a summary of activities we would like to complete at your institution during this study. We will work with your appointed coordinator to:

- Verify institution type and control, and enrollment information;
- Obtain enrollment lists for specified enrollment periods so that we can select a sample of students from these lists;
- Identify the best sources of locating information for all selected students and parents so that we can contact them for the student and parent surveys;
- Identify first-time first-year students so that we can obtain their SAT/ACT score or administer an assessment test;
- Identify sources for detailed demographic and academic information that we will collect for all sampled students; and
- Identify sources for financial aid information that we will collect for all sampled students who received aid for the 1989-90 school year.

These activities may differ somewhat according to your type of institution, your method of recordkeeping, and the degree to which you wish to rely upon Westat's field staff for collecting data under your supervision. Public reporting for your institution coordinator's part of this collection of information is estimated to vary from 8 to 40 hours per institution, with an average of 24 hours per institution over the duration of the study. This includes time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to the U.S. Department of Education; Information Management and Compliance Division, Washington, D.C. 20202-4651; and to the Office of Management and Budget, Paperwork Reduction Project 1850-0631, Washington, DC 20503.

We will work with your coordinator to devise a plan that will accomodate your preferences as much as possible while maintaining consistent and efficient data collection procedures.

an eligible student if so requested by the student's parents.

(20 U.S.C. 1232g(b)(1) and (b)(2)(A))

§ 99.31 Prior consent for disclosure not required.

(a) An educational agency or institution may disclose personally identifiable information from the education records of a student without the written consent of the parent of the student or the eligible student if the disclosure is—

(1) To other school officials, including teachers, within the educational institution or local educational agency who have been determined by the agency or institution to have legitimate educational interests;

(2) To officials of another school or school system in which the student seeks or intends to enroll, subject to the requirements set forth in § 99.34;

(3) Subject to the conditions set forth in § 99.35, to authorized representatives of:

(i) The Comptroller General of the United States,

(ii) The Secretary,

(iii) The Commissioner, the Director of the National Institute of Education, or the Assistant Secretary for Education, or

(iv) State educational authorities;

(4) In connection with financial aid for which a student has applied or which a student has received; *Provided*, That personally identifiable information from the education records of the student may be disclosed only as may be necessary for such purposes as:

(i) To determine the eligibility of the student for financial aid,

(ii) To determine the amount of the financial aid,

(iii) To determine the conditions which will be imposed regarding the financial aid, or

(iv) To enforce the terms or conditions of the financial aid;

(5) To State and local officials or authorities to whom information is specifically required to be reported or disclosed pursuant to State statute adopted prior to November 19, 1974. This subparagraph applies only to statutes which require that specific information be disclosed to State and local officials and does not apply to statutes which permit but do not require disclosure. Nothing in this paragraph shall prevent a State from further limiting the number or type of State or local officials to whom disclosures are made under this subparagraph;

(6) To organizations conducting studies for, or on behalf of, educational agencies or institutions for the purpose of developing, validating, or administering predictive tests, administering student aid programs, and improving instruction; *Provided*, That the studies are conducted in a manner which will not permit the personal identification of students and their parents by individuals other than representatives of the organization and the information will be destroyed when no longer needed for the purposes for which the study was conducted; the term "organizations" includes, but is not limited to, Federal, State and local agencies, and independent organizations;

(7) To accrediting organizations in order to carry out their accrediting functions;

(8) To parents of a dependent student, as defined in section 152 of the Internal Revenue Code of 1954;

(9) To comply with a judicial order or lawfully issued subpoena; *Provided*, That the educational agency or institution makes a reasonable effort to notify the parent of the student or the eligible student of the order or subpoena in advance of compliance therewith; and

(10) To appropriate parties in a health or safety emergency subject to the conditions set forth in § 99.36.

(b) This section shall not be construed to require or preclude disclosure of any personally identifiable information from the education records of a student by an educational agency or institution to the parties set forth in paragraph (a) of this section.

(20 U.S.C. 1232g(b)(1))

§ 99.32 Record of requests and disclosures required to be maintained.

(a) An educational agency or institution shall for each request for and each disclosure of personally identifiable information under § 99.31 with the understanding that the information will be redisclosed to other parties under that section; *Provided*, That the recordkeeping requirements of § 99.32 are met with respect to each of those parties.

(c) An educational agency or institution shall, except for the disclosure of directory information under § 99.37, inform the party to whom a disclosure is made of the requirement set forth in paragraph (a) of this section.

(20 U.S.C. 1232g(b)(4)(B))

records of a student, maintain a record kept with the education records of the student which indicates:

(1) The parties who have requested or obtained personally identifiable information from the education records of the student, and

(2) The legitimate interests these parties had in requesting or obtaining the information.

(b) Paragraph (a) of this section does not apply:

(i) To requests by or disclosure to a parent of a student or an eligible student;

(ii) To requests by or disclosures to school officials under § 99.31(a)(1);

(iii) If there is written consent of a parent of a student or an eligible student, or

(iv) To requests for or disclosure of directory information under § 99.37.

(c) The record of requests and disclosures may be inspected:

(1) By the parent of the student or the eligible student,

(2) By the school official and his or her assistants who are responsible for the custody of the records, and

(3) For the purpose of auditing the recordkeeping procedures of the educational agency or institution by the parties authorized in, and under the conditions set forth in § 99.31(a)(1) and (3).

(20 U.S.C. 1232g(b)(4)(A))

(41 FR 24670, June 17, 1976, as amended at 42 FR 4460, Jan. 25, 1977)

§ 99.33 Limitation on redisclosure.

(a) An educational agency or institution may disclose personally identifiable information from the education records of a student only on the condition that the party to whom the information is disclosed will not disclose the information to any other party without the prior written consent of the parent of the student or the eligible student, except that the personally identifiable information which is disclosed to an institution, agency or organization may be used by its officers, employees and agents, but only for the purposes for which the disclosure was made.

(b) Paragraph (a) of this section does not preclude an agency or institution from disclosing personally identi-

A-6

1990 NATIONAL POSTSECONDARY STUDENT AID STUDY

AUTHORIZATION AND CONFIDENTIALITY

The NPSAS was designed and is being conducted by the National Center for Education Statistics (NCES), U.S. Department of Education in compliance with the mandate stated in section 406(g) and of the General Education Provisions Act (GEPA), as amended, P.L. 100-297 (Hawkins-Stafford education amendments of 1988), subsection i:

> *As of March 30, 1990, and not less than every 3 years thereafter, the Center shall conduct a national study and survey of financial aid in accordance with the provisions of section 1303(c) of the Higher Education Amendments of 1986...Concurrent with each survey, the Center shall conduct longitudinal studies of freshmen and graduating students concerning access, choice, persistence, progress, curriculum and attainment...*

As part of this study NCES' contractor, Westat, Inc. of Rockville, Maryland has been authorized to collect information on selected students from institutional records. The grant of this authority is made pursuant to the provisions of the Family Education Rights and Privacy Act of 1976 (FERPA) (20 U.S.C. 1232g). FERPA allows the release of institutional record information to the Secretary of Education or his agent without prior consent of survey members, in connection with the "enforcement of Federal legal requirements" [20 U.S.C. 1232g (b) (3)]. Since the National Center for Education Statistics is an authorized representative of the Secretary of Education, compliance with NCES' official request for information constitutes compliance with the Federal Education Law (406) cited above.

P.L. 100-297 also established new confidentiality requirements for NCES. Requirements are that all records on individuals must be kept confidential by NCES and its contractors under penalty of law; disclosure is punishable by up to 5 years incarceration and/or up to a $250,000 fine. A plan for assuring the confidentiality of this project has been developed by NCES and Westat. This plan includes signed confidentiality agreements obtained from all personnel who will have access to individual identifiers, personnel training regarding the meaning of confidentiality, and controlled access to computer files.

A copy of the relevant sections of FERPA is shown on the reverse side for your review. Should you have any additional concerns, please contact one of the NCES Co-Project Officers, Dr. Sandra Garcia (202) 357-6597 or Mr. Gerald Malitz (202) 357-6364.

UNITED STATES DEPARTMENT OF EDUCATION

OFFICE OF THE ASSISTANT SECRETARY
FOR EDUCATIONAL RESEARCH AND IMPROVEMENT

February 1990

Dear Student,

You have been selected to participate in a study to determine how you and your family are financing your education. This study, the National Postsecondary Student Aid Study, is being conducted by Westat, a national research company located in the Washington, D.C. area, for the United States Department of Education. The results will be used to help determine future Federal policy regarding student financial aid.

Westat will contact you by phone in the near future in order to ask you some questions concerning your experiences in securing financial aid, your experiences with financial aid programs, and the effects financing your education has had on you and your family. Since during the course of the interview we will be asking some questions about your income and expenses, it may help to shorten the interview to have your financial records available when we call.

New guidelines put forth in the 1988 revisions to the Paperwork Reduction Act of 1980 (5 CFR Part 1320) require that the estimated number of burden hours must be displayed on each data collection instrument or relayed to the respondent in some way. Therefore, we are providing you with the following required statement.

> *Public reporting burden for this collection of information is estimated to vary from 20 to 40 minutes per response, with an average of 30 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to the U.S. Department of Education, Information Management and Compliance Division, Washington, D.C. 20202-4651; and to the Office of Management and Budget, Paperwork Reduction Project 1850-0631, Washington, D.C. 20503.*

We want to assure you that under Federal law all information obtained from the National Postsecondary Student Aid Study will be kept strictly confidential and cannot be disclosed or released to your school or any other group or individual. In fact:

> *All information which would permit identification of individuals will be regarded as strictly confidential, will be used only for purposes of the study, and will not be disclosed or released to any other group or individual. All information obtained by the study will be combined and reported only in statistical form. This study is authorized by law (20 U.S. C. 1221e.1). While you are not required to respond, your cooperation is needed to make the results of this survey comprehensive, accurate and timely.*

If you have any questions about the study prior to Westat calling you, or if you would like to set up an appointment to be interviewed, please call Westat's toll free number 1-800-937-8288.

Thank you very much for your cooperation.

Sincerely,

Emerson J. Elliott
Acting Commissioner

P.S. In order to ensure that you receive this letter, copies may have been sent to more than one address.

WASHINGTON, D.C. 20208

HOW WILL THE 1990 NPSAS BE CONDUCTED?

Between December 1989 and March 1990, interviewers visited 1,150 of the nation's 12,000 postsecondary institutions. These institutions represent all types of postsecondary schools: public, private, 2-year, 4-year, large, and small. Interviewers collected enrollment and financial aid data for 60,000 students. Similar information from 10,000 students who were enrolled in the 1989-90 school year (but not in the fall) will be collected. Beginning in the spring 1990, Westat, Inc., a national research company, will interview 70,000 students and 27,000 parents by telephone.

The 1990 NPSAS also includes a longitudinal component. It will study the possible effects of student aid, changes in family financial situations over time, progress through school, and changes in education and employment goals and aspirations. First-time students in 1990 NPSAS will provide base year data for a followup study to be conducted at two-year intervals, beginning in 1992.

All information collected during the study which would permit identification of individuals will be kept strictly confidential, will be used only for purposes of the study, and will not be disclosed or released to any other group or individual. Federal laws with penalties up to $250,000 and 5-year prison terms protect NPSAS data. All information obtained by the study will be combined and reported only in statistical form.

NEED MORE INFORMATION ON NPSAS?

Please contact Dr. Andrew G. Malizio, (202) 357-6448 or Dr. C. Dennis Carroll, (202) 357-6774, Longitudinal Studies Branch, National Center for Education Statistics, 555 New Jersey Avenue, NW, Washington, DC 20208-5652. Or contact the NPSAS Operations Office at Westat, Inc., (toll-free), at 1-(800) 937-8288.

NEED MORE INFORMATION OR QUESTIONS ABOUT STUDENT AID?

Please call the:

Federal Student Aid Information Center
(toll-free) 1-(800) 333-4636
(9:00 a.m. -- 5:30 p.m., Eastern Time)
Mon. -- Fri.

WHAT IS THE NATIONAL POSTSECONDARY STUDENT AID STUDY?

The National Postsecondary Student Aid Study (NPSAS) is the most comprehensive nationwide study of how students and their families pay for postsecondary education. It includes nationally representative samples of undergraduates, graduates, and first-professional students; students attending less than 2-year institutions, 2- to 3-year schools, 4-year colleges, and major universities. Students who receive financial aid, as well as those who do not receive aid, and their parents participate in NPSAS. Results of the study are used to help determine future Federal policy regarding student financial aid. The National Center for Education Statistics, U.S. Department of Education, conducts the study every three years.

The first NPSAS was conducted during the 1986-87 school year. Data were gathered from about 1,100 colleges, universities, and other postsecondary institutions; 43,000 students; and 13,000 parents. These data provided information on the cost of postsecondary education, the distribution of financial aid, and the characteristics of both aided and nonaided students and their families. The first of two major reports, *Undergraduate Financing of Postsecondary Education* and *Student Financing of Graduate and Professional Education*, has been published.

About 46 percent of the 11.2 million undergraduate students enrolled in the fall of 1986 received some type of financial aid. The average amount of aid awarded to full-time, full-year undergraduates was about $3,800. Amounts varied considerably depending on the type of schools that students attended.

About 57 percent of the 1.3 million graduate and first-professional students enrolled in the fall of 1986 received some type of financial aid. The average amount of aid awarded to full-time postbaccalaureate students enrolled was about $9,600. The amounts of aid awarded, however, varied considerably depending on the type of schools and students' programs.

WHAT IS NPSAS TRYING TO FIND OUT?

More than 13 million students were enrolled at postsecondary institutions in fall 1989. During the 1989-90 school year, more than $26 billion in financial aid (from Federal and State governments, and institutions) will be awarded to students to help meet their education expenses. Information collected as part of 1990 NPSAS will address such questions as:

- What are the costs to students and their families of obtaining postsecondary education?
- How do students and families pay for postsecondary education?
- Why do some students from apparently similar backgrounds receive more financial aid than other students?
- How much have students and families borrowed to pay for postsecondary education?
- What are the family characteristics of aided and nonaided students?
- How do aided and nonaided students compare on total resources available for education and other expenses?
- How do costs and financial aid influence students' choice of schools and majors?
- How have costs and financial aid packages changed since the 1986-87 school year?
- Why do some students and families not apply for financial aid?
- Do aided and nonaided students complete their programs?

1990 NPSAS

OMB# 1850-0631
Expires: 12/89

- Please review the name and address of your institution on the label below and make any necessary changes in the box.

INSTITUTION: _____

STREET: _____

CITY _____ STATE: _____ ZIP: _____

- Please provide the name and telephone number of the person we should contact about your participation in this study.

COORDINATOR NAME: _____

TITLE: _____ PHONE: (___) _____

- Please check the appropriate category in each box.

Control

- ☐ Public
- ☐ Private, nonprofit
- ☐ Private, for-profit

Highest Level Offered

- ☐ < 2-year only
- ☐ 2 or 3 year only
- ☐ Bachelor's
- ☐ Master's
- ☐ Doctor's
- ☐ First-Professional

- What is your estimated 1989-90 undergraduate enrollment?____

943521

BUSINESS REPLY MAIL

FIRST CLASS PERMIT NO. 1596 ROCKVILLE, MD

POSTAGE WILL BE PAID BY ADDRESSEE

npsas

National Postsecondary Student Aid Study

Post Office Box 10428
Rockville, Maryland 20850

NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES

Appendix B

Analysis Variables Specifications

APPENDIX B. Information on NPSAS:90 Derived Variables and Analysis File

The following variables are contained on the derived variable file on the restricted-access compact disk. For more information on how these variables were defined, users should refer to the electronic codebook, which is available on the compact disk. Contact the Statistical Standards and Methodology Division. Most of the variables on the derived variable file are available in the NPSAS:90 Table Generation Software, available while supplies last from NCES. Otherwise, the NPSAS:90 and the NPSAS:87 Table Generation systems (CD-ROM) are available [at $23 each] from the Superintendent of Documents, PO Box 371954, Pittsburgh, PA 15250-7954. NPSAS:87 is stock #065-000-00471-2; and NPSAS:90 is stock #065-00472-1.

```
/  NPSASID  'NPSAS student ID based on PSU sort'
/  PSKEEPWT 'Weight'
/  PSU      'Probability sampling unit'
/  CONTROL  'Institution control'
/  LEVEL    'Institution type'
/  OFCON1   'Institution type and control (summary)'
/  AFFILTN  'Institutional affiliation'
/  DATASRC  'Data sources available'
/  AGE      'Age as of 12/31/89'
/  MARITAL  'Marital status'
/  HSDEG    'High school degree or equivalent'
/  GENDER   'Gender of student'
/  LOCALRES 'Local residence'
/  CTZNSHP  'Citizenship'
/  RACE     'Race/ethnicity'
/  RACE2    'Race/ethnicity not incl. Hispanic'
/  HISPANIC 'Hispanic origin'
/  PARLOAN  'Total loans from parents'
/  APPLYNSH 'Number of schools student applied to'
/  ACCEPTAT 'Num of schls to which student was acceptd'
/  RELIGION 'Religion'
/  STUOCC1  'Occupation - primary spell'
/  STUIND1  'Industry - primary spell'
/  CHOICE   'Was sample school first choice'
/  OFERDFA1 'Financial aid importance'
/  DISTANCE 'Sample school distance from perm home'
/  MAJORS   'Major field or study'
/  AVEEXP   'Average monthly household expenses'
/  ACTVDUTY 'Currently on active duty'
/  VETERAN  'Veteran of U.S. armed forces'
/  DISABLTY 'Any disabilities'
/  COMMSERV 'Ever done community service'
/  COMSERHR 'Hrs per week doing community service'
/  PROGTYP  'Type of degree program'
/  UGRDLVL1 'Undergraduate level'
/  GPA      'Cumulative GPA'
/  GPACAT   'Cumulative GPA (categories)'
/  LENGTHCL 'Length of progrm for clock-hr students'
/  HRSPERWK 'Clock hrs required/week at sample schl'
/  CREDHRS  'Credit hours at sample school'
/  ATTEND   'FT/PT attendance status at sample schl'
/  NOENROLL 'Num of periods enrolled at sample schl'
/  NOSCH    'Number of schools attended'
/  SATV     'SAT score-verbal'
/  SATM     'SAT score-math'
/  SATTOTAL 'SAT score-combined'
/  ACT      'ACT score'
/  ATTNSTAT 'Attendance status'
/  COMPTO87 'Comparable to NPSAS 87 sample'
/  ENLEN    'Length of enrollment'
/  PSTSECYR 'Year first enrolled in postsec educ'
/  TOTCOST  'Total student costs'
/  TUITCOST 'Total tuition and fees'
```


BEST COPY AVAILABLE

```
/  ROOMCOST 'Room and board'
/  OTHRMCST 'Other room costs'
/  BOOKCOST 'Books and supplies'
/  OTHRCOST 'Other educational expenses'
/  OFFCOST  'Other off-campus costs'
/  INJURIS  'In jurisdiction for tuition'
/  PELLTUIT 'Pell budget: tuition and fees'
/  PELLROOM 'Pell budget: room, board, books, etc.'
/  PELLCHIL 'Pell budget: child care'
/  PELLHAND 'Pell budget: handicapped'
/  CMTUIT   'Congressional Methodology (CM) tuition and fees'
/  CMBOOKS  'CM books and supplies'
/  CMROOM   'CM room and board'
/  CMTRANS  'CM transportation'
/  CMMISC   'CM miscellaneous'
/  CMDPNDNT 'CM dependent'
/  CMHANDCP 'CM handicapped'
/  ENROLL88 'Institution enrollment in 1988'
/  CMCOSTS  'CM total costs'
/  CMBUDGET 'CM non-tuition/fees total costs'
/  CSTPERFC 'Ratio of total CM costs to EFC'
/  ANYCWS   'Received any Coll Work-Study'
/  CWSAMT   'Total College Work-Study (amount)'
/  SCHOLSHP 'Received scholarship/asstship'
/  SCHOLAMT 'Total scholarship/asstship (amount)'
/  WAIVER   'Tuition waivers and discounts'
/  WAIVAMT  'Total tuition waivers'
/  SPSEMP   'Spouse employed'
/  SPSINC   'Spouse's income'
/  EMPLPRD  'Employment period (summer, term, both)'
/  CWSPERND 'Total Coll Work-Study earned'
/  WKINC    'Total student earnings from work'
/  WKINCCAL 'Total student earnings from work in 1989'
/  EVRBORW  'Have borrowed for undergrad education'
/  BORAMT1  'Amount borrowed for undergrad education'
/  BORAMT2  'Amount borrowed for grad education'
/  STILLOWE 'Still owe money for education'
/  OWEAMT   'Amount student still owes'
/  STSAVPLN 'Used prepayment or savings plan'
/  SAVBONDS 'Used U.S. Savings Bonds'
/  EXEDCOL  'Highest level expected to complete'
/  PARMAR   'Parent's marital status'
/  REFPAR   'Referent parent'
/  REFCONTR 'Referent parent amt of contribution'
/  REFLOAN  'Referent parent amt of loan'
/  NREFCON  'Non-referent parent contribution'
/  NREFLOAN 'Non-referent parent amt of loan'
/  REFINC88 'Income of referent parent in 1988'
/  REFINC89 'Income of referent parent in 1989'
/  PARCONTR 'Total contribution from parents'
/  REFSIK   'Referent parent provided support-in-kind'
/  EDSAVING 'Referent parent used money from ed savings'
/  EDTRUST  'Referent parent used money from trust fund'
/  NOTFORED 'Referent parent used money from othr savings'
/  BORROW   'Referent parent borrowed (not mortgage)'
/  SECMORG  'Referent parent took out second mortgage'
/  RFINANC  'Referent parent refinanced any real estate'
/  MOREJOBS 'Referent parent took additional job'
/  MOREHRS  'Referent parent worked more hours per week'
/  CURINC   'Referent parent used regular job income'
/  SELLASET 'Referent parent sold assets'
/  RETFUNDS 'Referent parent used retirement funds'
/  OTHFUNDS 'Referent parent used any other funds'
/  PLUSLOAN 'Referent parent obtained a PLUS loan'
/  STATLOAN 'Referent parent got state-sponsored loan'
/  SCHLLOAN 'Referent parent got school-sponsored loan'
/  SIGNLOAN 'Referent parent obtained a signature loan'
/  HOMELOAN 'Referent parent obtained home equity loan'
```

```
/  CREDLOAN 'Referent parent obtained a line of credit'
/  LIFELOAN 'Referent parent loan against life insurance'
/  COMMLOAN 'Referent parent obtained commercial loan'
/  UNDRLOAN 'Undergraduate student level'
/  SMAELOAN 'Referent parent got Family Ed Ln (Sallie Mae)'
/  RETRLOAN 'Referent parent loan against retirement fund'
/  OTHRLOAN 'Referent parent obtained any other loan'
/  PREPAY   'Referent parent used tuition prepayment plan'
/  BONDPROG 'Referent parent in U.S. Ed Savings Bond program'
/  DADEDUC  'Father's highest level of education'
/  MOMEDUC  'Mother's highest level of education'
/  DADTRADE 'Father's trade school length'
/  MOMTRADE 'Mother's trade school length'
/  DADUNIV  'Father's amt of college education'
/  MOMUNIV  'Mother's amt of college education'
/  DEPEND   'Dependency status'
/  EFC3     'Expected family contribution (composite)'
/  SAI      'Student aid index'
/  NONFMCST 'CM cost-EFC'
/  PELLDIFF 'Pell max - SAI'
/  DEPINC   'Dependent student's 1988 family AGI'
/  INDEPINC 'Independent stud/spouse's 1988 AGI'
/  FAMINC   '1988 AGI'
/  FAMFARM  'Family farm'
/  FARMVAL  'Value of family farm'
/  APPFORM  'Primary application used'
/  FEDTAXES 'Federal taxes paid'
/  UNTAXINC 'Untaxed income'
/  FAMNUM   'Number in family'
/  POSTED   'Number in postsecondary education'
/  OTHERTAX 'Allowance for state/other taxes'
/  AIDSRC1  'Title IV based source of financial aid'
/  AIDSRC2  'Source of financial aid'
/  AIDPACK  'Type of financial aid package'
/  FEDPACK  'Type of Federal aid package'
/  PELLPACK 'Type of aid package containing Pell'
/  STAFPACK 'Type of aid pack containing Stafford'
/  LOANPACK 'Type of loan package'
/  MNSTAT1  'July 1989 enroll/employ status'
/  MNSTAT2  'August 1989 enroll/employ status'
/  MNSTAT3  'September 1989 enroll/employ status'
/  MNSTAT4  'October 1989 enroll/employ status'
/  MNSTAT5  'November 1989 enroll/employ status'
/  MNSTAT6  'December 1989 enroll/employ status'
/  MNSTAT7  'January 1990 enroll/employ status'
/  MNSTAT8  'February 1990 enroll/employ status'
/  MNSTAT9  'March 1990 enroll/employ status'
/  MNSTAT10 'April 1990 enroll/employ status'
/  MNSTAT11 'May 1990 enroll/employ status'
/  MNSTAT12 'June 1990 enroll/employ status'
/  IPEDS2   'IPEDS ID for second school attended'
/  IPEDS3   'IPEDS ID for third school attended'
/  IPEDS4   'IPEDS ID for fourth school attended'
/  ANYAID   'Student receive any aid'
/  OTHSCAID 'Student receive aid at other schools'
/  PELL     'Student receive Pell Grant'
/  SEOG     'Student receive SEOG grant'
/  CWSP     'Student receive CWSP award'
/  PERKINS  'Student receive Perkins Loan'
/  STAFFORD 'Student receive Stafford Loan'
/  PLUS     'Student receive PLUS loan'
/  SLS      'Student receive SLS loan'
/  ICL      'Student receive ICL loan'
/  OTHERFED 'Student receive other Federal aid'
/  TITLEIV  'Student receive Title IV aid'
/  FEDAID   'Student receive any Federal aid'
/  T4LOAN1  'Student rec any Title IV ln but PLUS'
/  T4LOAN2  'Student rec any Title IV ln inc PLUS'
```

```
/  FEDLOAN1 'Student rec any Fed loan but VA/DOD'
/  FEDLOAN2 'Student rec any Fed loan inc VA/DOD'
/  CAMPUS   'Student receive campus based aid'
/  INSTITUT 'Student rec any institutional aid'
/  INCWSFL  'Student receive institutional CWS'
/  INNEEDFL 'Student rec inst Need Based aid'
/  INNONDFL 'Student rec inst non-Need Based aid'
/  INNONDF  'Student rec inst non-Need Based grant'
/  INNEDF   'Student rec inst Need Based grant'
/  STATE    'Student receive any state aid'
/  STNEEDFL 'Student rec state Need Based aid'
/  STNONDFL 'Student rec state non-Need Based aid'
/  SSIG     'Student receive SSIG'
/  OTHER    'Student receive any other aid'
/  TEACHAST 'Student receive teaching assistship'
/  RESAST   'Student receive research assistship'
/  ASTSHP   'Student receive assistantship-type unkn'
/  FELLSHP  'Student receive fellowship'
/  TRNSHP   'Student receive traineeship'
/  OTHGRAD  'Student receive other grad award'
/  TOTGRNT  'Student receive any grant'
/  TOTLN    'Student receive any loan'
/  TOTWK    'Student receive any work study'
/  OTHERCAT 'Student rec other aid (not grt,ln,CWS)'
/  NFEDAID  'Student receive non-Federal aid'
/  FEDGRT   'Student receive Federal grant'
/  NFEDGRT  'Student receive non-Federal grant'
/  FEDLN    'Student receive Federal loan'
/  NFEDLN   'Student receive non-Federal loan'
/  FEDOTH   'Student receive Federal other'
/  NFEDOTH  'Student receive non-Federal other'
/  TOTAID   'Total aid'
/  OTHSCAMT 'Aid amount at non-sampled schools'
/  PELLAMT  'Federal aid:  Pell Grant (amount)'
/  SEOGAMT  'Federal aid:  SEOG grant (amount)'
/  CWSPAMT  'Federal aid:  CWSP awarded (amount)'
/  PERKAMT  'Federal aid:  Perkins Loan (amount)'
/  STAFFAMT 'Federal aid:  Stafford Loan (amt)'
/  PLUSAMT  'Federal aid:  PLUS loan (amount)'
/  SLSAMT   'Federal aid:  SLS loan (amount)'
/  ICLAMT   'Federal aid:  ICL loan (amount)'
/  OTHFDAMT 'Federal aid:  Other (amount)'
/  TITIVAMT 'Amount of Title IV aid'
/  TFEDAID  'Total Federal aid'
/  T4AMT1   'Amt of Title IV loans except PLUS'
/  T4AMT2   'Amt of Title IV loans include PLUS'
/  FEDAMT1  'Amount of Fed loans except VA/DOD'
/  FEDAMT2  'Amount of Fed loans include VA/DOD'
/  CAMPAMT  'Amount of campus based aid'
/  INSTAMT  'Amount of institutional aid'
/  INSTCWS  'Institutional aid:  CWS (amount)'
/  INSTNEED 'Institutional aid: Need Based (amt)'
/  INSTNOND 'Inst aid: non-Need Based (amt)'
/  INNONDGR 'Inst aid: non-Need Based grant (amt)'
/  INNEEDGR 'Inst aid: Need Based grant (amount)'
/  STATEAMT 'Amount of state aid'
/  STATNEED 'Amount of Need Based state aid'
/  STATNOND 'Amount of non-Need Based state aid'
/  SSIGAMT  'Amount of SSIG'
/  OTHERAID 'Amt of othr aid (not Fed/state/inst)'
/  TEACHAMT 'Teaching assistantship amount'
/  RESAMT   'Research assistantship amount'
/  ASTAMT   'Assistantship amount'
/  FELLAMT  'Fellowship amount'
/  TRNAMT   'Traineeship amount'.
/  OTHGRAMT 'Other graduate amount'
/  TOTGRT   'Amount of grant aid'
/  TOTLOAN  'Amount of loan aid'
```


BEST COPY AVAILABLE

```
/  TOTWKST  'Amount of work study aid'
/  TOTOTHR  'Amt of othr aid (not grant/ln/CWS)'
/  TNFEDAID 'Total non-Federal aid'
/  TFEDGRT  'Total Federal grants'
/  TNFEDGRT 'Total non-Federal grants'
/  TFEDLN   'Total Federal loans (except PLUS)'
/  TNFEDLN  'Total non-Federal loans'
/  TFEDOTHR 'Total Federal other (including PLUS)'
/  TNFEDOTH 'Total non-Federal other'
/  T4PK1    'Student receive Pell,Stafford, or CWSP'
/  T4PK1AMT 'Amount of Federal Pell,CWSP, & Staff'
/  ASSIST   'Student receive any assistantship'
/  ASSTAMT  'Assistantship amount (all types)'
/  EFC1     'Expected family contribution (reported)'
/  EFC2     'Expected family contribution (derived)'
/  EMPLYAMT 'Amount of aid from employer'
/  EMPLOYER 'Student receive aid from employer'
/  FATHEDUC 'Father's education'
/  INCOME   'Dependency and income level'
/  IPEDSID  'IPEDS ID for sampled school'
/  MAJRCODE 'Major or program of study'
/  MOTHEDUC 'Mother's education'
/  OFCONL   'Institution type & level'
/  PAREDUC  'Parent's education'
/  PSEYR1   'Year first enrolled in postsec educ'
/  STUIND   'Industry - primary spell'
/  STUOCC   'Occupation - primary spell'
/  TYPAGE   'Typical age'
/  T4PACK1  'Pell, CWSPERND, or Stafford'.
```

EXAMPLES OF OTHER VARIABLES ON THE RESTRICTED-ACCESS CD-ROM

The following information was collected from those students who reported during the telephone interview that their expenses were greater than their available resources (including financial aid).

```
/  APPLOAN  'Did student apply for loans/take more loans'
/  ADDJOB   'Did student work or take additional job'
/  ASKPARNT 'Did student ask parents for money/more money'
/  REDUCELD 'Did student reduce course load'
/  CUTDOWN  'Did student cut down on expenses'
/  WITHDRAW 'Did student withdraw from school'
/  TRANSFER 'Did student transfer to cheaper school'
/  BACKHOME 'Did student move back home'
/  REMEDY   'Take other action to help w/expenses'
/  WLDYOUDO 'What other action did student take'
```

The following information was collected from those students who did not apply for financial aid.

```
/  FAMPAY   'Family and student could pay for education'
/  NODEBT   'Student not willing to go into debt for school'
/  HIINCOME 'Family income too high to qualify'
/  LOWGRADE 'Grades/test scores too low to qualify'
/  HARDAPP  'Too hard to apply for aid'
/  NODISCLO 'Didn't wish disclose financial situation'
/  NOELIGBL 'Ineligible-attended school part-time'
/  NOAIDMON 'No money was available for aid'
/  MISDLINE 'Missed deadline for application'
/  IMPORTNT 'Most important reason student did not apply'
```

The following information was collected from those students who refused at least a portion of financial aid. (Reasons why students ever refused financial aid)

```
/  GOINDEBT 'Loans offered, didn't want to go in debt'
/  NOTNEED  'Student did not need assistance'
/  INTERFER 'Work-study offered, interfered with schl'
/  EARNEMPL 'Work-study less than earned at other job'
/  OTRREFUS 'Other reason aid was refused'
```

The following information was collected to specify health disabilities.

```
   /  DEAFNESS 'Is student hearing impaired or deaf'
   /  SPEECH   'Does student have speech limitation/disability'
   /  ORTHO    'Does student have orthopedic disability/limit'
   /  LEARNDIS 'Does student have learning disability'
   /  VISUAL   'Is student partially sighted or blind'
   /  HEALTOTH 'Student have other health related disability'.
```

The following information was collected from first-time students only.

```
   /  CHOICE   'Is student attending first choice institution'
   /  OFERDFA1 'Was financial aid important in decision'
   /  STUDYPL  'Student/family have specific place for study'
   /  DALYNEWS 'Student/family have a daily newspaper'
   /  REFBOOKS 'Student/family have encyclopedia/ref books'
   /  TYPEWRIT 'Student/family have a typewriter'
   /  DISHWASH 'Student/family have a dishwasher'
   /  TWOCARS  'Student/family have two or more vehicles'
   /  BOOKS50  'Student/family have more than 50 books'
   /  OWNROOM  'Student has room of own at home'
   /  CALCUL   'Student/family have a pocket calculator'
   /  VCR      'Student/family have a VCR'
   /  COMPUTER 'Student/family have a personal computer'
   /  MWORKPS  'Mother work before student went to elem school'
   /  MWORKPST 'Mother work full/part-time before elem'
   /  MWORKEL  'Mother work in elementary school years'
   /  MWORKELT 'Mother work full/part-time elem schl yrs'
   /  MWORKHS  'Mother work in high school years'
   /  MWORKHST 'Mother work full/part-time in hs years'
   /  REMREAD  'Number of hours of remedial reading'
   /  REMWRITE 'Number of hours of remedial writing'
   /  REMMATH  'Number of hours of remedial math'
   /  REMSTSK  'Number of hrs of remedial study skills'
   /  COMPAREA 'Academic ability compared to others'
   /  COMPAREB 'Artistic ability compared to others'
   /  COMPAREC 'Drive to achieve compared to others'
   /  COMPARED 'Emotional health compared to others'
   /  COMPAREE 'Leadership ability compared to others'
   /  COMPAREF 'Mathematical ability compared to others'
   /  COMPAREG 'Physical health compared to others'
   /  COMPAREH 'Popularity compared to others'
   /  COMPAREI 'Intellectual confidence compared to othr'
   /  COMPAREJ 'Social self-confidence compared to othrs'
   /  COMPAREK 'Writing ability compared to others'
   /  COMPAREL 'Mechanical or technical ability compared'.
==========================================================================
```

APPENDIX B.2 Analysis Variables Specifications

This section of Appendix B contains (in alphabetical order by variable name) the specifications used to create the NCES derived variables. In the case of award amounts, these specifications assume that the amounts have first been aggregated to the full-year amount. Source flags are not listed separately, but all use the following scheme:

0=Missing, no imputation performed
1=Record Abstract, not adjusted
2=Record Abstract, adjusted
3=Student Survey, not adjusted
4=Student Survey, adjusted
5=Parent Survey, not adjusted
6=Parent Survey, adjusted
7=More than one source, not adjusted
8=More than one source, adjusted
9=Imputed from other variables

Award flags primarily are defined as dichotomous variables, with "1" indicating that aid (of the specified type) was received, and "2" indicating that no aid was received. However, five variables (PELL, SEOG, STAFFORD, PLUS, and SLS) used a different scheme:

1=No aid received
2=Some aid received
3=Maximum received.

List of variables (excluding award flags and source flags)

(NOTE: SS refers to student telephone interview; RA refers to the institution data collected on the student Record Abstract; PS refers to the parent telephone interview. ABCODE refers to the Award/Budget code from the Award/Budget data module; which was based on information from the Record Abstract. The numbers in parentheses refer to the item number during the telephone interview. Users interested in the item wording, or additional information on specific variable definitions should request a copy of the NPSAS:90 Electronic Codebook from NCES. This codebook describes codes, labels, frequencies, item wording, and software to produce fully-labeled SPSS and SAS code.)

ACCEPTAT=SS(9.02)

ACT=RA(32b)

ACTVDUTY=SS(6.15)
If CITIZEN=2 then ACTVDUTY=-1

AGE=89-SS(06b)
If missing, use RA(2)
If still missing, check Pell recipient file
If still missing, check RA(73_24, 74_1, 75_7, and 76_4)

ANYCWS=TOTWK (duplicate)

APPFORM=RA(72)

APPLYNSH=SS(9.01)
If APPLYNSH=50 then APPLYNSH=APPLYNSH/10

```
ASTAMT=RA(54e), assistantships only [ABCODE=651=659 and F_ASSIST=1]
      If (RESAMT+TEACHAMT+ASTAMT)<SS(3.11a) then ASTAMT=SS(3.11a) for
            sample school
      If at non-sampled school, ASTAMT=sum(ASTAMT,SS(3.11a) for non-
sampled schools)

ATTEND
      If RA(33) ne 2 or RA(34e) missing, then ATTEND=RA(35d)
      If missing then ATTEND=RA(34e)
      If missing, then ATTEND=SS(2.12a or 2.13a) (based on sampled term)
      If missing, take from Pell file

ATTNSTAT
If ENLEN < 9 then                    (part year)
      If ATTEND=1 and for every term SS(2.12a or 2.12b)=1 then
ATTNSTAT=3
      If ATTEND=2 or 3 or for some term SS(2.12a or 2.12b)=2 or 3 then
ATTNSTAT=6
If SS(1.04)=2 then                          (one school)
      If ATTEND=1 and for every term SS(2.12a or 2.12b)=1 then
ATTNSTAT=1
      If ATTEND=2 or 3 or for some term SS(2.12a or 2.12b)=2 or 3 then
ATTNSTAT=4
If SS(1.04)=1 then                          (more than 1)
      If ATTEND=1 and for every term SS(2.12a or 2.12b)=1 then
ATTNSTAT=2
      If ATTEND=2 or 3 or for some term SS(2.12a or 2.12b)=2 or 3 then
ATTNSTAT=5

AVEEXP=sum of SS(2.17) components
      =RENTMORT+FOOD+TRANSP+PERSONAL+KIDCARE+EDLOANS+OTHEREXP

BONDPROG=PS(1.15)

BOOKCOST=SS(2.12e-f--summed over all terms)

BORAMT1=SS(3.16a)
      If BORAMT1=0 then BORAMT1=-1

BORAMT2=SS(3.16b)
      If BORAMT2=0 then BORAMT2=-1

BORROW=PS(1.07d)

CAMPAMT=SEOGAMT + CWSPERND + PERKAMT

CHOICE=SS(9.03)

CMBOOKS=RA(70b_1)
      If missing, then CMBOOKS=RA(65b_1)

CMBUDGET=CMCOSTS-CMTUIT

CMCOSTS=CMTUIT+CMBOOKS+CMROOM+CMTRANS+CMMISC+CMDPNDNT
            +CMHANDCP

CMDPNDNT=RA(70f_1)
      If missing, then CMDPNDNT=RA(65f_1)

CMEFC=non-missing component of RA(70h) and sum of RA(70h_1,h_2,h_3) for
sample term
```

```
      If missing, then use RA(65h) and RA(65h_1,h_2,h_3)

CMHANDCP=RA(70G_1)
      If missing, then CMHANDCP=RA(65g_1)
      If ATTEND=3 then CMHANDCP=0

CMMISC=RA(70e_1)
      If missing, then CMMISC=RA(65e_1)
      If ATTEND=3 then CMMISC=0

CMROOM=RA(70c_1)
      If missing, then CMROOM=RA(65c_1)
      If ATTEND=3 then CMROOM=0
If LOCALRES=3 and ATTEND ne 3 and [RA(73_4) ne 1 and RA(74_14) ne 1 and
RA(75_15d) and RA(76_18d) ne 1 and SS(8.04a) ne 1] and CMROOM<1500 then
CMROOM=1500
Else if LOCALRES ne 1 and ATTEND ne 3 and CMROOM<2500 then CMROOM=2500

CMTRANS=RA(70d_1)
      If missing, then CMTRANS=RA(65d_1)

CMTUIT=RA(70a_1)
      If missing, then CMTUIT=RA(65a_1)

COMMLOAN=PS(1.08h)

COMMSERV=SS(6.25)

COMPTO87 -- Comparable to NPSAS:87 sample
      If enrolled in fall '89 and not in Puerto Rico then COMPTO87=1
      Else COMPTO87=2

COMSERHR=SS(6.28)
      If COMSERV=2 then COMSERHR=0

CONTROL (from sampling)
1 = Public
2 = Private
3 = Proprietary
CREDHRS
      If RA(35b2)=2, 3, or 4 then CREDHRS=RA(35b1)
      If RA(35b2)=1 then CREDHRS=RA(35b1) x 2/3
      If missing then
            If SS(2.12c1b)=1 then CREDHRS=SS(2.12c1a)
            If SS(2.12c1b)=2 then CREDHRS=SS(2.12c1a) x 2/3
CREDLOAN=PS(1.08f)

CSTPERFC=CMCOSTS/EFC

CTZNSHP
If SS(6.12)=1 then CTZNSHP=1
      Else if SS(6.13)=1 then CTZNSHP=2
      Else if SS(6.13)=2 then CTZNSHP=3
If missing, CTZNSHP=RA(25)
If still missing, use Pell recipient file
If still missing, CTZNSHP=RA(76_12)
If still missing, then
      if RA(73_25, or 74_4, or 75_8a)=1 then CTZNSHP=1
      else if RA(73_25, or 74_4, or 75_8a)=2 then
            if FEDAID=1 then CTZNSHP=2
            else CTZNSHP=3
```

```
CURINC=PS(1.07i)

CWSAMT=TOTWKST (duplicate)

CWSPAMT=RA(39c)  [ABCODE=103]
      If not at sample school, then CWSPAMT=CWSPAMT+SS(3.08a)
      If CWSPAMT>TOTCOST>500 then do
            If sum of SS(3.08a) for all schools (including sample
school) <= TOTCOST
                  and sum of SS(3.08a)>0 then CWSPAMT=sum of SS(3.08a)
            Else CWSPAMT=TOTCOST

CWSPERND=RUT(2a1b)  [ABCODE=104]
      If missing or not at sample school, then
            if SS(3.08b2)><1 [STAWSAMT] and SS(3.08c2)<1 [USRWSAMT] then
                  use WSTUDINC from student file
      If still missing, use CWSPAMT
Range edits for CWSPAMT and CWSPERND
      If CWSPAMT=CWSPERND and CWSPAMT>7500 then
            CWSPAMT=CWSPAMT/10; CWSPERND=CWSPERND/10;
      Else if CWSPAMT>7500
            DIFFCWSP=CWSPERND-CWSPAMT
            CWSPAMT=CWSPAMT/10
            CWSPERND=CWSPAMT+DIFFCWSP

DADEDUC=SS(7.02_1)

DADTRADE=SS(7.02a1)

DADUNIV=SS(7.02b1)

DATASRC -- data sources available
      1=Record Abstract, Student Survey, and Parent Survey
      2=Record Abstract and Student Survey only
      3=Record Abstract only
      4=Student Survey and Parent Survey only
      5=Student Survey only
```

```
DEPEND (Dependency status)
If RA(57a)=2 then depend=2 (institutional judgment)
  Else if AGE > 23 then depend=2 (age)
  Else if RA(76_18c, or 75_15c, or 74_13, or 73_3)=1 then depend=2
(orphan)
      If RA info missing, use SS(8.01)
  Else if MARITAL=2-3 and RA(76_20,or 74_24,or 73_5,or 75_21)=2 then
depend=2 (marital)
      If RA info missing, use SS(8.02c)
  Else if RA(75_15d, or 73_4, or 74_14, or 76_18d)=1 then depend=2
(legal dependents)
      If RA info missing, use SS(8.04a)
  Else if RA(73_2, 74_12, 75_15b, 76_18b)=1 then depend=2 (veteran)
      If RA info missing, use SS(6.16)
  Else if PROGTYP=5-8 and RA(76_20,or 74_24,or 73_5,or 75_21)=2 then
depend=2 (grad)
      If RA info missing, use SS(8.02c)
  Else if PROGTYP=1-4 and MARITAL=1 and RA(73_[12 or 13] and 73_14, or
74_[21 or 22] and 74_23, or 75_[19b or 20a] and 75_20b)=1 and RA(74_15
and 74_16, or 73_6 and 73_7, or 75_16a and 75_16b, or 76_19a and
76_19b)=2 then depend=2 (undergrad)
      If RA info missing, use SS(8.03c and d) in place of 1st set,
SS(8.02a and b) in place of 2nd
  Else if RA(57a)=1 then depend=1;
Else if AGE not missing and
[RA(76_18c, or 75_15c, or 74_13, or 73_3)=2 or [SS(8.01)=2] and
[RA(75_15d, or 73_4, or 74_14, or 76_18d)=2 or SS(8.04a)=2] and
[RA(73_2, 74_12, 75_15b, 76_18b)=2 or SS(6.16)=2] then do;
If (PROGTYP=5-8 or MARITAL=2 or MARITAL=3) and [RA(76_20,or 74_24,or
73_5,or 75_21)=1 or SS(8.02c)=1] then depend=1;
Else if (PROGTYP=1-4, or 9 and MARITAL=1) and [[RA(73_[12 or 13] or
73_14, or 74_[21 or 22] or 74_23, or 75_[19b or 20a] or 75_20b)=2 or
SS(8.03c or 8.03d)=2] or [RA(74_15 or 74_16, or 73_6 or 73_7, or 75_16a
or 75_16b, or 76_19a or 76_19b)=1 or SS(8.02a or 8.02b)=1]] then
depend=1;
End;
  If missing, take value from update (U_DEPEND)
  If still missing, take value from Pell recipient file
  If still missing, and 20<=AGE<=23 and (5<=PROGTYP<=8 or MARITAL=2)
      then DEPEND=2
  If still missing, and MARITAL=2 then DEPEND=2
  If still missing, and 0<AGE<=23 then DEPEND=1
  If still missing, and PROGTYP>4 then DEPEND=2

DEPINC (defined only if DEPEND=1; otherwise equals -1)
      DEPINC=RA(73_76)
      If missing, then DEPINC=RA(74_36b)
      If missing, then DEPINC=RA(75_35)
      If missing, then DEPINC=RA(76_1a)
      If missing, use Pell recipient file
      If missing, then DEPINC=PS(3.13)
      If missing, use SS(7.04) REFINC88
      If missing, then impute

DISABLTY=SS(6.24)

DISTANCE=SS(2.02)

EDSAVING=PS(1.07a)

EDTRUST=PS(1.07b)
```

EFC1=non-missing component of RA(70h) and sum of RA(70h_1,h_2,h_3) for sample term
If missing, then use RA(65h) and RA(65h_1,h_2,h_3)
If missing, use Pell recipient file

EFC2=expected family contribution as derived by formula

EFC3=EFC1
If DEPEND=1 and -1<EFC1<700 and (UGRDLVL1=1 or UGRDLVL1=2) and UGRDLVL2<3 then EFC3=700
Else if DEPEND=1 and -1<EFC1<900 then EFC3=900
Else if DEPEND=2 and -1<EFC1<1200 and LEGALDEP=2 then EFC3=1200
If EFC1<0 then EFC3=EFC2
(NOTE: EFC3 is the variable used in the NPSAS:90 Table Generation System)

EMPLYANT
If EMPLYAMT=0 and SA311B>0 then EMPLYAMT=EMPLYAMT+SA311B
If not at sampled school, EMPLYAMT=EMPLYAMT+SS311B

EMPLPRD -- Employment period
If EMSTAT1-EMSTAT12= 2 for every month (no employment at any time) then EMPLPRD=1
If MNSTAT1-MNSTAT12 has at least one (2 or 4) and at least one 5, but no 1 or 3
(employment, not in school terms)
then EMPLPRD=2
If MNSTAT1-MNSTAT12 has at least one (1 or 3), but no 5
(employment in school terms only)
then EMPLPRD=3
If MNSTAT1-MNSTAT12 has at least one (1 or 3), and at least one 5
(employment both in and out of school)
then EMPLPRD=4

EMSTAT1-EMSTST12 -- monthly employment status (for each month separately)
1=Employed at some time in the month
2=Not employed in any of the month

ENLEN
Count total number of months from 7/89 to 6/90 that student was enrolled based on SS(1.07)
e.g., ENLEN=0
If 7/89 ge mnthb/yearb and le mnthe/yeare for some term then ENLEN=ENLEN+1
If 8/89 ge mnthb/yearb and le mnthe/yeare for some term then ENLEN=ENLEN+1
And so on through 6/90

ENSTAT1-ENSTST12 -- monthly enrollment status (for each month separately)
1=Enrolled full time at some time in the month
2=Enrolled part time at some time in the month
3=Not enrolled in any of the month

EVRBORW
If SS(3.16)>0 then EVRBORW=1
If SS(3.16)=0 then EVRBORW=2
If missing and PROGTYP=1-4 and TOTLOAN>0 then EVRBORW=1
If missing and PROGTYP=1-4 and TOTLOAN=0 then EVRBORW=2

(Note: this last is really an imputation; it assumes that if the student isn't currently borrowing any money, he/she didn't borrow any money earlier. The data would be biased if we allow for TOTLOAN>0 but not TOTLOAN=0.)

```
FAMFARM
      If DEPEND=2 then do;    /* student */
            FAMFARM=RA(73_63)
            If FAMFARM<0 then FAMFARM=RA(75_80)
            If FAMFARM<0 and SS(8.12e)>0 then FAMFARM=1;
                  Else if SS(8.12e)=0 then FAMFARM=2;
                  [SS variable is ASETFARM]
            end;
      If DEPEND=1 then do;    /*  parents */
            FAMFARM=RA(73_102)
            If FAMFARM<0 then FAMFARM=RA(75_57)
            If FAMFARM<0 and PS(3.09d)>0 then FAMFARM=1
                  Else if PS(3.09d)=0 then FAMFARM=2;
                  [PS variable is WRTHFARM]
            end;
      If FAMFARM<0 and RA(74_57)>0 then FAMFARM=RA(74_57)
      If FAMFARM=3 then FAMFARM=-9

FAMINC
      If DEPEND=1 then FAMINC=DEPINC
      If DEPEND=2 then FAMINC=INDEPINC

FAMNUM
      If DEPEND=2 then do;    /* student */
            FAMNUM=RA(73_30)
            If FAMNUM<0 then FAMNUM=RA(74_25)
            If FAMNUM<0 then FAMNUM=RA(75_22)
            If FAMNUM<0 then FAMNUM=SS(8.04b);
                  [SS variable is RDEPENDS]
            If FAMNUM<0 and SS(8.04b)=-1 and MARITAL=1 then FAMNUM=1;
            If FAMNUM<0 and SS(8.04b)=-1 and MARITAL=2 and SS(8.04a)=2
                  then FAMNUM=2;
                  [SS(8.04a) is LEGALDEP]
                        end;
      If DEPEND=1 then do;    /*  parents */
            FAMNUM=RA(73_72)
            If FAMNUM<0 then FAMNUM=RA(74_27)
            If FAMNUM<0 then FAMNUM=RA(75_29)
            If FAMNUM<0 then FAMNUM=RA(76P2)
            If FAMNUM<0 then FAMNUM=PS(7.05)
                  [PS variable is TOTSUPP]
            end;

FARMVAL  /* this makes an assumption that farm value=business value */
      If DEPEND=2 then do;    /* student */
            FARMVAL=SS(8.12e);
                  [SS variable is ASETFARM]
            If FARMVAL<0 and FAMFARM=1 then FARMVAL=RA(73_61)
            If FARMVAL<0 and FAMFARM=1 then FARMVAL=RA(75_78)
            end;
      If DEPEND=1 then do;    /*  parents */
            FARMVAL=PS(3.09d)
                  [PS variable is WRTHFARM]
            If FARMVAL<0 and FAMFARM=1 then FARMVAL=RA(73_100)
            If FARMVAL<0 and FAMFARM=1 then FARMVAL=RA(75_55)
            end;
```

```
      If FARMVAL<0 and RA(74_57)=1 and DEPEND=2 then FARMVAL=RA(74_55a);
      If FARMVAL<0 and RA(74_57)=1 and DEPEND=1 then FARMVAL=RA(74_55b);
      If FARMVAL=0 and FAMFARM=2 then FARMVAL=-1;

FATHEDUC
      FATHEDUC=-9
      If DADEDUC=1 or DADEDUC=2 or DADEDUC=3 then FATHEDUC=DADEDUC
      If DADTRADE>0 then FATHEDUC=DADTRADE+3
      If FATHEDUC<0 and DADEDUC=4 then FATHEDUC=4
      If DADUNIV>0 then FATHEDUC=DADUNIV+6
      If FATHEDUC<0 and DADEDUC=5 then FATHEDUC=7


FEDAMT1=T4AMT1+RA(39i)+RA(39j)+RA(39m)+(RA(39n) if TYPE=2)
  =T4AMT1+[ABCODE=110]+[ABCODE=111]+[ABCODE=114]+[ABCODE=151-199 if
TYPE=2]
      If school not sample school, then FEDAMT1=FEDAMT1+SS(3.06b-d)

FEDAMT2=FEDAMT1+RA(48h)+(RA(48i-k) if TYPE=2)
  =FEDAMT1+[ABCODE=408]+[ABCODE=451-499 if TYPE=2]

FEDTAXES
      If DEPEND=2 then do;   /* student */
            FEDTAXES=RA(73_35)
            If FEDTAXES<0 then FEDTAXES=RA(74_37a)
            If FEDTAXES<0 then FEDTAXES=RA(75_61)
            If FEDTAXES<0 then FEDTAXES=RA(76_4b)
            If FEDTAXES<0 then FEDTAXES=SS(8.08b);
                  [SS variable is INCTAX88]
            end;
      If DEPEND=1 then do;   /*  parents */
            FEDTAXES=RA(73_77)
            If FEDTAXES<0 then FEDTAXES=RA(74_37b)
            If FEDTAXES<0 then FEDTAXES=RA(75_36)
            If FEDTAXES<0 then FEDTAXES=RA(76_4a)
            end;

FELLAMT=RA(54c)  [ABCODE=603]
      For sample school, if FELLAMT<SS(3.10a-d) then FELLAMT=SS(3.10a-d)
      If school not sample school, then FELLAMT=FELLAMT+SS(3.10a-d)

FORMSA=RA(71a)

FORMSB=RA(71b)

FORMSC=RA(71c)

FORMSD=RA(71d)

FORMSE=RA(71e)

FORMSF=RA(71f)

FORMS7=RA(71g)

GENDER=SS(6.06a)
      If missing, use RA(5)
      In some cases, might impute based on CIP codes (majors)

GPA (use cumulative GPA first; use most recent GPA if cumulative not
available)
```

```
      If RA(38)=1 then GPA=RA(37)
      If RA(38)=2 then GPA=RA(37)-1
      If RA(38)=3 then GPA=RA(37) - 5.5
            If GPA>4.0 then GPA=4.0
            If GPA<0 then GPA=0
      If RA(38)=4 then GPA=0.1xRA(37) - 5.5
            If GPA>4.0 then GPA=4.0
            If GPA<0 then GPA=0
      If GPA<0 and UGRDLVL1=1 and RA(37c)=1 then GPA=-1

GPACAT
      If 0<=GPA<1 then GPACAT=1
      If 1.0<=GPA<2 then GPACAT=2
      If 2.0<=GPA<3 then GPACAT=3
      If 3.0<=GPA then GPACAT=4
      Else GPACAT=-9

HOMELOAN=PS(1.08e)

HRSPERWK=RA(34c)

HSDEG=SS(6.20)
      If missing, use RA(22)

ICLAMT=RA(39h)  [ABCODE=109]

If schools not in campus-based program (NOSCH=1 and WESID=11011063,
1371277, 1492131, or 3161106) then shift funds in CWSPERND, PERKAMT, and
SEOGAMT to OTHFDAMT
and set CWSPAMT, CWSPERND, PERKAMT, and SEOGAMT to 0

INDEPINC (defined only if DEPEND=2; otherwise equals -1)
      INDEPINC=RA(73_34)
      If missing, then INDEPINC=RA(74_36a)
      If missing, then INDEPINC=RA(75_60)
      If missing, then INDEPINC=RA(76_1b)
      If missing, use Pell recipient file
      If missing, then INDEPINC=SS(8.07b)
      If missing, then impute

INJURIS -- In jurisdiction for tuition purposes
INJURIS=RA(36)

INNEEDGR=RA(45f) [ABCODE=306]

INNONDGR=RA(45e)  [ABCODE=305]

INSTAMT=RA(45a-k)  [ABCODE=401-499]
If RA(39a)>100 and PELLAMT=0 then INSTAMT=INSTAMT+RA(39A)
If SS(3.04a3)>100 and PELLAMT=0 then INSTAMT=INSTAMT+SS(3.04A3)
If RA(39a) > PELLAMT> 100 then INSTAMT=INSTAMT+RA(39a)-PELLAMT
If school not sample school then
     INSTAMT=INSTAMT + SS(3.05c-d) + SS(3.06f) + SS(3.10c) + SS(3.15)
Note:  there is additional information in SS(3.08b), but we can't
distinguish state from institutional aid.

INSTNEED=RA(45f)+(RA(45i-k) if KIND=5 or 7)
=[ABCODE=306]+[ABCODE=351-399 if KIND=5 or 7]
If RA(39a)>100 and PELLAMT=0 then INSTNEED=INSTNEED+RA(39A)
If SS(3.04a3)>100 and PELLAMT=0 then INSTNEED=INSTNEED+SS(3.04A3)
If RA(39a) > PELLAMT> 100 then INSTNEED=INSTNEED+RA(39a)-PELLAMT
```

INSTNOND=RA(45e)+(RA(45i-k) if KIND=6)
=[ABCODE=305]+[ABCODE=351-399 if KIND=6]

IPEDS1-IPEDS3
Take FICE codes developed from list in SS(1.05) [SCHLNAME]

IPEDS1-IPEDS3 -- IPEDS ID for each non-sampled school at which enrolled
Note: the record abstract often did not collect sufficient information to properly identify the school

LENGTHCL=RA(34b)
If missing and RA(33)=2 then LENGTHCL=SS(2.12c2 or 2.13c2)
(depending which term is sampled term)

LEVEL (from sampling)
1 = Less than 2 yrs
2 = 2-3 years
3 = 4 years not PHD
4 = 4 years PHD

LIFELOAN=PS(1.08g)

MAJRCODE=RA(35n) [CIP code]
If missing, then MAJRCODE=RA(34n) [CIP code]
If missing, then MAJRCODE= [CIP code from SS(2.06 and 2.06)

MARITAL
If SS(4.01)=2, 4, or 5 then MARITAL=1
If SS(4.01)=1 then MARITAL=2
If SS(4.01)=3 then MARITAL=3
If missing, then MARITAL=RA(21)
If still missing, then use Q73S27, Q74S5, or S76S12 (RA)
MNSTAT1-MNSTST12 -- monthly status (for each month separately)
1=Enrolled full-time and employed
2=Enrolled full-time and not employed
3=Enrolled part-time and employed
4=Enrolled part-time and not employed
5=Not enrolled and employed
6=Not enrolled and not employed

MOMEDUC=SS(7.02_1)

MOMTRADE=SS(7.02a1)

MOMUNIV=SS(7.02b1)

MOREHRS=PS(1.07h)

MOREJOBS=PS(1.07g)

MOTHEDUC
MOTHEDUC=-9
If MOMEDUC=1 or MOMEDUC=2 or MOMEDUC=3 then
MOTHEDUC=MOMEDUC
If MOMTRADE>0 then MOTHEDUC=MOMTRADE+3
If MOTHEDUC<0 and MOMEDUC=4 then MOTHEDUC=4
If MOMUNIV>0 then MOTHEDUC=MOMUNIV+6
If MOTHEDUC<0 and MOMEDUC=5 then MOTHEDUC=7

NOENROLL -- Number of periods at sample school

```
      Count terms at sample school from SS(1.07)
            [Already calculated]
      If missing, count number of terms in record update (1 to 4)

NONFMCST=CMCOSTS - CMEFC

NOSCH -- Number of schools attended
      Count schools in SS(1.05) and add 1 (for sample school)

NOTFORED=PS(1.07c)

NPSASID -- sequential student ID
      Starts at 1 for undergraduates
      Starts at 70,001 for graduate students and first professionals

NREFCON
      If REFPAR=1 or 5 then NREFCON=SS(4.06a1)     [MOMPAY]
      If REFPAR=2 or 6 then NREFCON=SS(4.06a2)     [DADPAY]
      If REFPAR=3 or 4 then NREFCON=-1;
      If NREFCON<-1 then NREFCON=-9
NREFLOAN
      If REFPAR=1 or 5 then NREFLOAN=SS(4.06b1)     [MOMLOAN]
      If REFPAR=2 or 6 then NREFLOAN=SS(4.06b2)     [DADLOAN]
      If REFPAR=3 or 4 then NREFLOAN=-1;
      If NREFLOAN<-1 then NREFLOAN=-9

OFCON1
If CONTROL=1 then do (public)
      If level=1 then OFCON=01 (Pub <2-years)
      If level=2 then OFCON=02 (Pub 2-3 years)
      If level=3 then OFCON=03 (Pub 4-yr no PHD)
      If level=4 then OFCON=04 (Pub 4-yr PHD)
(continued. . )
If CONTROL=2 then do (private, not-for-profit)
      If level=1 then OFCON=05 (Priv <2-years)
      If level=2 then OFCON=06 (Priv 2-3 years)
      If level=3 then OFCON=07 (Priv 4-yr no PHD)
      If level=4 then OFCON=08 (Priv 4-yr PHD)
If CONTROL=3 then do (proprietary)
      If level=1 then OFCON=09 (Prop <2-years)
      If level=2-4 then OFCON=10 (Prop 2+ years)

OFCON2
      If OFCON1=01 or 02 then OFCON2=OFCON1
      If OFCON1=03 or 04 then OFCON2=03
      If OFCON1=05 or 06 then OFCON2=04
      If OFCON1=07 or 08 then OFCON2=05
      If OFCON1=09 or 10 then OFCON2=06

OFCONL
      OFCONL=OFCON2
      If PROGTYP>4.5 and CONTROL=1 then ONFONL=7
      If PROGTYP>4.5 and CONTROL=2 then ONFONL=8
      If PROGTYP>4.5 and CONTROL=3 then ONFONL=8

OFERDFA1=SS(9.04)

OFFCOST=SS(2.17c-g) x (ENLEN - # of months for which ROOMCOST is
defined)
      If OFFCOST>12000 then OFFCOST=OFFCOST/12
```

```
OTHERAID=RA(48a-k, 51a-j)  [ABCODE=401-499,501-599]
If at sampled school,
      If RA(51a)=0 then OTHERAID=OTHERAID+SS(3.11b)
      If RA(51d)=0 then OTHERAID=OTHERAID+SS(3.05e1b)+SS(3.05e1c)
      If RA(51e)=0 then OTHERAID=OTHERAID+SS(3.05e1d)
      If OTHERAID<(SS(3.06g)+SS(3.05e)+SS(3.11b)) then
            OTHERAID=SS(3.06g)+SS(3.05e)+SS(3.11b)
If not at sampled school,
OTHERAID=OTHERAID+SS(3.06g)+SS(3.05e)+SS(3.11b)

OTHERTAX
      If DEPEND=2 then OTHERTAX=RA(76_6b);   /* student */
      If DEPEND=1 then OTHERTAX=RA(76_6a);   /*  parents */

OTHFDAMT=RA(39i-n)  [ABCODE=110-114,151-199]
      If school not sample school, then OTHFDAMT=OTHFDAMT+SS(3.06b-d)

OTHFUNDS=PS(1.071)

OTHGRAMT=RA(54e), but not assistantships  [ABCODE=651-699 and F_ASSIST
ne 1]

OTHRCOST=SS(2.12g-h--summed over all terms)

OTHRLOAN=PS(1.08OTH)

OTHRMCST=SS(2.17a-b) x (ENLEN - # of months for which ROOMCOST is
defined)
      If OTHRMCST>24000 then OTHRMCST=OTHRMCST/12

OTHSCAMT=SS(3.03a)  (if at non-sampled school)

OWEAMT=SS(3.17)
      If BORAMT1>0 and BORAMT2>0 and OWEAMT=-1 then OWEAMT=0

PAREDUC
      PAREDUC=FATHEDUC
      If MOTHEDUC>PAREDUC then PAREDUC=MOTHEDUC

PARCONTR=REFCONTR+NREFCON
      If REFCONTR=-1 then PARCONTR=NREFCON
      If NREFCON=-1 then PARCONTR=REFCONTR
      IF REFCONTR=-1 and NREFCON=-1 then PARCONTR=0
      If REFCONTR=-9 or NREFCON=-9 then PARCONTR=-9

PARLOAN=REFLOAN+NREFLOAN
      If REFLOAN=-1 then PARLOAN=NREFLOAN
      If NREFLOAN=-1 then PARLOAN=REFLOAN
      IF REFLOAN=-1 and NREFLOAN=-1 then PARLOAN=0
      If REFLOAN=-9 or NREFLOAN=-9 then PARLOAN=-9

PELLAMT=(amount from Pell recipient file)
            If MATCHFLG=3 then PELLAMT=0
      =0 if not on Pell recipient file
      If PELLAMT>2300 then PELLAMT=PELLAMT/2
      If 0<PELLAMT<100 then PELLAMT=PELLAMT*10

PELLCHIL=RA(61c_1)
      If PELLCHIL>1000 then PELLCHIL=1000

PELLDIFF=2300-SAI
```

```
PELLHAND=RA(61d_1)
      If PELLHAND>5000 then PELLHAND=5000

PELLROOM=RA(61b_1)

PELLTUIT=RA(61a_1)

PERKAMT=RA(39d)  [ABCODE=105]
      If school not sample school, then PERKAMT=PERKAMT+SS(3.04b5)
      If UGRDLVL=1 or 2 and PERKAMT>4500 then PERKAMT=4500
      If PERKAMT>TOTCOST>500 then PERKAMT=TOTCOST
      If PERKAMT>0 and NOSCH=1 and WESID=2271245 or 4141305 then
            OTHFDAMT=OTHFDAMT+PERKAMT
            PERKAMT=0
      If 0<PERKAMT<100 then PERKAMT=PERKAMT*10

PLUSAMT=RA(39f)  [ABCODE=107]
      If PLUSAMT=0 and PS(1.08aov)>0 then PLUSAMT=PS(1.08aov)
      If DEPEND=2 then PLUSAMT=0
      If PLUSAMT>4000 then PLUSAMT=4000
      If 0<PLUSAMT<100 then PLUSAMT=PLUSAMT*10
            If PLUSAMT still < 100 then PLUSAMT=PLUSAMT*10

PLUSLOAN=PS(1.08a)

POSTED
      If DEPEND=2 then do;   /* student */
            POSTED=RA(73_31)
            If POSTED<0 then POSTED=RA(74_26)
            If POSTED<0 then POSTED=RA(75_23)
            If POSTED<0 then POSTED=SS(8.04d);
                  [SS variable is NUMINCOL]
            If POSTED<0 and SS(8.04c=2) then POSTED=0;
                  [SS variable is ANYINCOL]
            end;
(continued . . )

      If DEPEND=1 then do;   /*  parents */
            POSTED=RA(73_73)
            If POSTED<0 then POSTED=RA(74_28)
            If POSTED<0 then POSTED=RA(75_30)
            If POSTED<0 then POSTED=RA(76P3)
            If POSTED<0 then POSTED=PS(7.06)
                  [PS variable is COLLSUPP]
            If FAMNUM=0 then POSTED=-1;
            end;

PREPAY=PS(1.14)

PROGTYP
      If RA(29)=1-8 then PROGTYP=RA(29)
      Otherwise, use SS(2.09) for sampled term
            If SS(2.09)=2 then PROGTYP=1 (Associate Degree)
            If SS(2.09)=3 then PROGTYP=2 (Bachelor^s Degr)
            If SS(2.09)=1 then PROGTYP=3 (Undrgrd Certifct)
            If SS(2.09)=5 then PROGTYP=5 (Master^s Degree)
            If SS(2.09)=6 then PROGTYP=6 (Doctoral Degree)
            If SS(2.09)=7 then PROGTYP=7 (First-Professnl)
            If SS(2.09)=4 or 92 then PROGTYP=8 (Other Grad Prgrm)
            If SS(2.09)=91 then PROGTYP=4 (Other undergrad)
      If still missing, use SS(2.08) for sampled term
```

```
            If SS(2.08)=6 then PROGTYP=5 (Master^s Degree)
            If SS(2.08)=7 then PROGTYP=6 (Doctoral Degree)
            If SS(2.08)=8 then PROGTYP=7 (First-Professnl)
      If still missing, use sampling information

PSTSECYR
      If SS(6.21)=89 then find earliest beginning year in SS(1.07)
            (It will be either 89 or 90)
      If SS(6.22a)=2 [POSTSEC] then
            If SS(2.01)=1 then find earliest beginning year in SS(1.07)
                  (It will be either 89 or 90)
      Else PSTSECYR=SS(6.22b_yr) [EDSTARTY]
      If missing then
            If UGRDLVL1=1 then find earliest beginning year in SS(1.07)
            If UGRDLVL1=2-5 then PSTSECYR=90-UGRDLVL1
      If PSTSECYR<20 then PSTSECYR=-9
RACE
If SS(6.07a)=3 then RACE=1 (American Indian)
If SS(6.07a)=4 then RACE=2 (Asian)
If SS(6.07a)=2 and SS(6.08) ne 1 then RACE=3 (Black, non-Hispanic)
If SS(6.07a)=1 or 2 and SS(6.08)=1 then RACE=4 (Hispanic)
If SS(6.07a)=1 and SS(6.08) ne 1 then RACE=5 (White, non-Hispanic)
If missing, then:
If RA(23)=3 then RACE=1 (American Indian)
If RA(23)=4 then RACE=2 (Asian)
If RA(23)=2 and RA(24)=2 then RACE=3 (Black, non-Hispanic)
If RA(23) ne 3 and 4 and [RA(24)=1 or SS(6.08)=1] then RACE=4 (Hispanic)
If RA(23)=1 and RA(24)=2 then RACE=5 (White, non-Hispanic)
If missing and SS(6.08) ne 1 and RA(23) ne 1, then:
      If SS(6.07a)=1 or RA(23)=1 then RACE=5
      Else if SS(6.07a)=2 or RA(23)=2 then RACE=3
If missing and student in Puerto Rico, RACE=4
If missing, then: (do these sequentially, dropping out as soon as an
assignment is made)
      If PS(4.01a or 4.01b)=3 then RACE=1
      If PS(4.01a or 4.01b)=4 then RACE=2
      If PS(4.03a or 4.03b)=1 then RACE=4
      If PS(4.01a or 4.01b)=2 then RACE=3
      If PS(4.01a or 4.01b)=1 then RACE=5
If missing, then impute

RACE2
If RACE<4 then RACE2=RACE
Else if RACE>4 then RACE2=RACE-1
Perform the following if RACE=4
If SS(6.07a)=3 then RACE2=1 (American Indian)
If SS(6.07a)=4 then RACE2=2 (Asian)
If SS(6.07a)=2 then RACE2=3 (Black)
If SS(6.07a)=1 then RACE2=4 (White)
If missing, then:
If RA(23)=3 then RACE2=1 (American Indian)
If RA(23)=4 then RACE2=2 (Asian)
If RA(23)=2 then RACE2=3 (Black)
If RA(23)=1 then RACE2=4 (White)
If missing, then: (do these sequentially, dropping out as soon as an
assignment is made)
      If PS(4.01a or 4.01b)=3 then RACE2=1
      If PS(4.01a or 4.01b)=4 then RACE2=2
      If PS(4.01a or 4.01b)=2 then RACE2=3
      If PS(4.01a or 4.01b)=1 then RACE2=4
If missing, then impute
```

```
REFCONTR
      If REFPAR=3 then do;
            REFCONTR=SS(4.05a)   [PARNTPAY]
            If REFCONTR=-1 then REFCONTR=0
            end;
      If REFPAR=1 or 5 then do;
            REFCONTR=SS(4.06a2)    [DADPAY]
            If REFCONTR=-1 then REFCONTR=0
            end;
      If REFPAR=2 or 6 then do;
            REFCONTR=SS(4.06a1)    [MOMPAY]
            If REFCONTR=-1 then REFCONTR=0
            end;
      If REFPAR=4 then REFCONTR=-1;
      If REFCONTR<-1 then REFCONTR=PS(1.03)     [AMTGIVE]
      If REFCONTR<-1 then REFCONTR=-9

REFINC88=(copy from SS)

REFINC89=(copy from SS)

PARMAR=PS(1.01)

REFLOAN
      If REFPAR=3 then REFLOAN=SS(4.05b)   [LOANPAR]
      If REFPAR=1 or 5 then REFLOAN=SS(4.06b2)    [DADLOAN]
      If REFPAR=2 or 6 then REFLOAN=SS(4.06b1)    [MOMLOAN]
      If REFPAR=4 then REFLOAN=-1;
      If REFLOAN<-1 then REFLOAN=PS(1.04)     [AMTLOAN]
      If REFLOAN<-1 then REFLOAN=-9

REFPAR=REFPAREN (from SS)

REFSIK=PS(1.05OV)

RELIGION=SS(9.08)

RESAMT=RA(54b)  [ABCODE=602]
      If (RESAMT+TEACHAMT+[RA(54e) and F_ASSIST=1])<SS(3.11a) for sample

            school, then RESAMT=0

RETFUNDS=PS(1.07k)

RETRLOAN=PS(1.08k)

RFINANC=PS(1.07f)

ROOMCOST=SS(2.14a,2.15--summed over all terms)

SAI
      Primary source:  Pell recipient file
      Secondary source:  RA(61e)
      Impute for all Pell recipients with missing SAI

SATM=RA(32a-math)

SATTOTAL=SATV+SATM
      If SATV<0 or SATM<0 then SATTOTAL=-9;

SATV=RA(32a-verbal)
```

```
SAVBONDS=SS(4.09b)

SCHLLOAN=PS(1.08c)

SCHOLAMT=TOTGRT+RA(54a,b,d)
      If school not sample school, then SCHOLAMT=SCHOLAMT+SS(3.11a)
[TEACHAMT]

SCHOLSHP
      If SCHOLAMT>0 then SCHOLSHP=1; else SCHOLSHP=2

SECMORG=PS(1.07e)

SELLASET=PS(1.07j)

SEOGAMT=RA(39b)  [ABCODE=102]
      If PROGTYP=5-8 then SEOGAMT=0
      If PROGTYP=1-4 and SEOGAMT>4000 then SEOGAMT=4000
      If PROGTYP=1-4 and ATTNSTAT=1 or 2 and 0<SEOGAMT<100 then
SEOGAMT=100
      If SEOGAMT>TOTCOST>500 then SEOGAMT=TOTCOST

SIGNLOAN=PS(1.08d)

SLSAMT=RA(39g)  [ABCODE=108]
      If school not sample school, then SLSAMT=SLSAMT+SS(3.06a)
      If SLSAMT>4000 then SLSAMT=4000
      If 0<SLSAMT<100 then SLSAMT=SLSAMT*10
            If SLSAMT still < 100 then SLSAMT=SLSAMT*10

SMAELOAN=PS(1.08j)

SPSEMP=SS(5.12a)

SPSINC=SS(5.12b)
STAFFAMT=RA(39e)  [ABCODE=106]
      If school not sample school, then STAFFAMT=STAFFAMT+SS(3.04b3)
      If UGRDLVL1=1 or 2 and UGRDLVL2<3 and STAFFAMT>2625 then
            STAFFAMT=2625
If PROGTYP=1-4 and UGRDLVL2 le 5 and (UGRDLVL1 ge 3 or UGRDLVL2 ge 3
            or UGRDLVL1=-9) and STAFFAMT>4000 then STAFFAMT=4000
      If PROGTYP=5-8 or UGRDLVL2>5 and STAFFAMT>7500 then
            STAFFAMT=7500
      If ATTEND=3 and SS(2.12a and 2.13a)=3 for every term then
STAFFAMT=0
      If 0<STAFFAMT<100 then STAFFAMT=STAFFAMT*10
            If STAFFAMT still < 100 then STAFFAMT=STAFFAMT*10

STATEAMT=RA(42a-j)  [ABCODE=201-299]
If school not sample school then STATEAMT=SS(3.05b)+SS(3.06e)+SS(3.10b)
Note:  there is additional information in SS(3.08b), but we can't
distinguish state from institutional aid.

STATLOAN=PS(1.08b)

STATNEED=RA(42d,42g)+(RA(42i-j) if KIND=5 or 7)
      =[ABCODE=204,207]+[ABCODE=251-299 if KIND=5 or 7]

STATNOND=RA(42e)+(RA(42i-j) if KIND=6)
      =[ABCODE=205]+[ABCODE=251-299 if KIND=6]
```

STILLOWE
If OWEAMT>0 then STILLOWE=1
Else if OWEAMT=0 then STILLOWE=2
Else STILLOWE=-9

STSAVPLN=SS(4.09a)

STUIND=SS(5.08) [coded]

STUOCC=SS(5.08) [coded]

T4AMT1=PERKAMT+STAFFAMT+SLSAMT+ICLAMT

T4AMT2=PERKAMT+STAFFAMT+PLUSAMT+SLSAMT+ICLAMT

T4PK1AMT=sum(PELLAMT,CWSPERND,STAFFAMT)

TEACHAMT=RA(54a) *[ABCODE=601]*
If (RESAMT+TEACHAMT+[RA(54e) and F_ASSIST=1])<SS(3.11a) for sample school, then TEACHAMT=0

TFEDAID=TITIVAMT+OTHFDAMT

TFEDGRT=PELLAMT + SEOGAMT + RA(39k) + RA(39l) + (RA(39n) if TYPE=1)
=PELLAMT + SEOGAMT + *[ABCODE=112]* + *[ABCODE=113]* + *[ABCODE=151-199 if TYPE=1]*
If not at sampled school, TFEDGRT=TFEDGRT+SS(3.10a)+SS(3.05a)

TFEDLN=PERKAMT+STAFFAMT+SLSAMT+ICLAMT+RA(39i,j,m) + (RA(39n) if TYPE=2)
=PERKAMT + STAFFAMT + SLSAMT + ICLAMT + *[ABCODE=110,111,114]* +
[ABCODE=151-199 if TYPE=2]
If not at sampled school, TFEDLN=TFEDLN+SS(3.06b-d)

TFEDOTHR=TFEDAID-TFEDGRT-TFEDLN-CWSPERND

TITIVAMT=PELLAMT+SEOGAMT+CWSPERND+PERKAMT+STAFFAMT+PLUSAMT+SLSAMT+ICLAMT

TNFEDAID=TOTAID-TFEDAID

TNFEDGRT=TOTGRT-TFEDGRT

TNFEDLN=TOTLOAN-TFEDLN

TNFEDOTH=OTHERAID-TFEDOTHR

TOTAID=TFEDAID+STATEAMT+INSTAMT+OTHERAID
Compare with SS(3.03a) summed over all schools to verify that no aid is left out

TOTCOST=TUITCOST+ROOMCOST+BOOKCOST+OTHRCOST+OTHRMCST+OFFCOST

TOTGRT=PELLAMT + SEOGAMT + RA(39k,l) + (RA(39n) if TYPE=1) + RA(42c-f) + (RA(42i-j) if TYPE=1) + RA(45a,c-g) + (RA(45i-k) if TYPE=1) + RA(48a-d,f,g) + (RA(48i-k) if TYPE=1) + RA(51a,c-e) + (RA(51h-j) if TYPE=1) + RA(54c) + (RA(54e) if TYPE=1)
=PELLAMT + SEOGAMT + *[ABCODE=112,113]* + *[ABCODE=151-199 if TYPE=1]* + *[ABCODE=203-206]* + *[ABCODE=251-299 if TYPE=1]* + *[ABCODE=301,303-307]* + *[ABCODE=351-399 if TYPE=1]* + *[ABCODE=401-404,406-407]* + *[ABCODE=451-499*

BEST COPY AVAILABLE

```
if TYPE=1] + [ABCODE=501,503-505] + [ABCODE=551-599 if TYPE=1] +
[ABCODE=603] + [ABCODE=651-699 if TYPE=1]
If not at sampled school,
      TOTGRT=TOTGRT+SS(3.10a-d)+SS(3.05a-e)+SS(3.11b)+SS(3.12a-b)
If at sampled school, add in SS(3.11b) if RA(51a=0)
If RA(39a)>100 and PELLAMT=0 then TOTGRT=TOTGRT+RA(39A)
If SS(3.04a3)>100 and PELLAMT=0 then TOTGRT=TOTGRT+SS(3.04A3)
If RA(39a) > PELLAMT> 100 then TOTGRT=TOTGRT+RA(39a)-PELLAMT

TOTLOAN=PERKAMT + STAFFAMT + SLSAMT + ICLAMT + RA(39i,j,m) + (RA(39n) if
TYPE=2) + RA(42g,h) + (RA(42i-j) if TYPE=2) + RA(45h) + (RA(45i-k) if
TYPE=2) + RA(48h) + (RA(48i-k) if TYPE=2) + RA(51f) + (RA(51h-j) if
TYPE=2) + (RA(54e) if TYPE=2)
=PERKAMT + STAFFAMT + SLSAMT + ICLAMT + [ABCODE=110,111,114] +
[ABCODE=151-199 if TYPE=2] + [ABCODE=207,208] + [ABCODE=251-299 if
TYPE=2] + [ABCODE=308] + [ABCODE=351-399 if TYPE=2] + [ABCODE=408] +
[ABCODE=451-499 if TYPE=2] + [ABCODE=506] + [ABCODE=551-599 if TYPE=2] +
[ABCODE=651-699 if TYPE=2]
If not at sampled school, TOTLOAN=TOTLOAN+SS(3.06a-g)

TOTOTHR=TOTAID-TOTGRT-TOTLOAN-TOTWKST

TOTWKST=CWSPERND+INSTCWS +RA(42b)=CWSPERND+INSTCWS+[ABCODE=202]
If not at sampled school, TOTWKST=TOTWKST+SS(3.08a-c)

TRNAMT=RA(54d)  [ABCODE=604]

TUITCOST [below definitions for RA(34d,35c) use total tuition and fees;
sum if needed]
      If RA(33) ne 2 or RA(34d) missing, then TUITCOST=RA(35c)
      If missing then TUITCOST=RA(34d)
      If other terms on SS, then TUITCOST=TUITCOST+SS(2.12d) for other
terms
      If missing then TUITCOST=SS(2.12d--summed over all terms)
      If still missing, get from IPEDS IC file
      If TUITCOST>25000 then TUITCOST=25000

UGRDLVL1 (undergraduate level for sampled term)
      If RA(30)=1-5 then UGRDLVL1=RA(30)
            Else if SS(2.08)=1-5 for sampled term, then
UGRDLVL1=SS(2.08)
      If missing, use Pell file
            If EDLEVEL=1 or 2 then UGRDLVL1=EDLEVEL
            If EDLEVEL=3 then imputation should be over range 3-5

UGRDLVL2 (other undergraduate level)
      Must look at individual terms (before combined)
      If SS(2.08)>8 then SS(2.08)=-1
      UGRDLVL2=max(SS(2.08) across terms)
      If UGRDLVL2 <= UGRDLVL1 then UGRDLVL2=-1

UNDRLOAN=PS(1.08i)

UNTAXINC
      If DEPEND=2 then do;   /* student */
            UNTAXINC=sum(RA(73_38-41)
            If UNTAXINC<0 then UNTAXINC=sum(RA(74_41a-44a))
            If UNTAXINC<0 then UNTAXINC=sum(RA(75_65-68))
            If UNTAXINC<0 then UNTAXINC=RA(76_2b)
            If UNTAXINC<0 then UNTAXINC=sum(SS(8.10a-d))
```

```
                [SS variables are SOCSEC88, AFDC88, KIDSUP88,
OTHINC88]
          end;
     If DEPEND=1 then do;   /*  parents */
          UNTAXINC=sum(RA(73_80-83)
          If UNTAXINC<0 then UNTAXINC=sum(RA(74_41b-44b))
          If UNTAXINC<0 then UNTAXINC=sum(RA(75_40-43))
          If UNTAXINC<0 then UNTAXINC=RA(76_2a)
          If UNTAXINC<0 then UNTAXINC=sum(PS(3.15a-e))
                [PS variables are SSI88, ADC88, CSUPP88, FDSTMP88,
OINCOM88]
          end;

VETERAN=SS(6.16)
     If ACTVDUTY=1 or CITIZEN=2 then VETERAN=-1

WAIVAMT=RA(45c)+RA(45d)
     If school not sample school then WAIVAMT=WAIVAMT+SS(3.15)

WAIVER
     If WAIVAMT>0 then WAIVER=1; else WAIVER=2

WKINC
     If SS(5.09)=-1 then SS(5.09)=0              /* ANYINCOM */
     If SS(5.10)=-1 then SS(5.10)=0              /* ESTINCOM */
     WKINC=SS(5.09)+SS(5.10)
     If SS(5.09)<-1 or SS(5.10)<-1 then WKINC=-9

WKINCCAL=SS(5.13)        /* TOTERN */
```

Appendix Table B.1.--Data sources available for students included in the analysis file, by student level and by dependency and receipt of student aid:

UGRAD Undergraduates only - NPSAS90 by DATASRC Data sources available

UGRAD	Count / Row Pct / Col Pct / Tot Pct	DATASRC: Abs, stud, prnt 1	Abs & stu survy 2	Abstract only 3	Std & pnt only 4	Std survey only 5	Row Total
1.00 DEP/AIDED		5888	6096	2858	17	11	14870
		39.6	41.0	19.2	.1	.1	31.9
		36.8	27.6	34.6	16.5	6.8	
		12.6	13.1	6.1	.0	.0	
2.00 DEP/NONAIDED		7078	2492	1111	43	24	10748
		65.9	23.2	10.3	.4	.2	23.1
		44.3	11.3	13.4	41.7	14.8	
		15.2	5.3	2.4	.1	.1	
3.00 IND/AIDED		2300	6868	3695	20	44	12927
		17.8	53.1	28.6	.2	.3	27.7
		14.4	31.1	44.7	19.4	27.2	
		4.9	14.7	7.9	.0	.1	
4.00 IND/NONAIDED		723	6636	608	23	83	8073
		9.0	82.2	7.5	.3	1.0	17.3
		4.5	30.0	7.4	22.3	51.2	
		1.6	14.2	1.3	.0	.2	
Column Total		15989	22092	8272	103	162	46618
		34.3	47.4	17.7	.2	.3	100.0

Number of Missing Observations: 170 (Information to determine dependentcy status was not available for 170 undergraduates).

GRAD Graduates only - NPSAS90 by DATASRC Data sources available

GRAD	Count / Row Pct / Col Pct / Tot Pct	DATASRC: Abs, stud, prnt 1	Abs & stu survy 2	Abstract only 3	Std survey only 5	Row Total
1.00 DEP/AIDED		3	400	80	1	484
		.6	82.6	16.5	.2	3.4
		30.0	3.1	6.3	1.7	
		.0	2.8	.6	.0	
2.00 DEP/NONAIDED		3	125	2		130
		2.3	96.2	1.5		.9
		30.0	1.0	.2		
		.0	.9	.0		
3.00 IND/AIDED		1	6488	1147	12	7648
		.0	84.8	15.0	.2	53.4
		10.0	49.9	90.8	20.7	
		.0	45.3	8.0	.1	
4.00 IND/NONAIDED		3	5988	34	45	6070
		.0	98.6	.6	.7	42.4
		30.0	46.1	2.7	77.6	
		.0	41.8	.2	.3	
Column Total		10	13001	1263	58	14332
		.1	90.7	8.8	.4	100.0

Number of Missing Observations: 0

DEP=Dependent IND=Independent
Data sources 1=Abstract, Student interview, and Parent interview
2=Abstract and Student interview
3=Abstract only
4=Student and Parent interviews only
5=Student survey only

Appendix C

Procedures Used for Data Imputations

Variable Name: HSDEG
Description: High School Degree or Equivalent
Cases Eligible for Imputation: All cases without valid values

Number with valid responses:	57,134
Number eligible for imputation:	3,986
Number not eligible for imputation:	N/A
Total:	61,120

Pre-Analysis:

Cross-tabulations were run on the following variables to determine the best predictors: OFCON1 PROGTYPC (undergraduate, graduate, and first professional) RACE GENDER GPACAT MARITAL

Hotdeck Imputation:

Sort by:

1. OFCON1
*2. PROGTYPC
3. RACE
4. MARITAL
5. SCHOOL

* No break.

Number imputed by Hotdeck imputation = 3,986

Post Analysis:

High School Degree or Equivalent

Value	Label	Before Imputation		After Imputation	
		Frequency	Percent	Frequency	Percent
1	Diploma	53,548	93.7	57,203	93.6
2	GED	1,946	3.4	2,090	3.4
3	Certificate	377	0.7	422	0.7
4	No H.S. Degree	1,263	2.2	1,405	2.3
-9	Not specified	3,986	--	--	--
Total		61,120	100.0	61,120	100.0

Variable Name: UGRDLVL1
Description: Undergraduate Level
Cases Eligible for Imputation: All cases without valid values

Number with valid responses:	44,072
Number eligible for imputation:	2,716
Number not eligible for imputation:	N/A
Total:	46,788

Pre-Analysis:

Cross-tabulations were run on the following variables to determine the best predictors:

OFCON1, GPAYES (1 = GPA available, 2 = GPA not available)

Hotdeck Imputation:

Sort variables:

*1. OFCON1
2. GPAYES
3. AGE
4. SCHOOL

*No break

Number imputed by Hotdeck imputation = 2,716

Post Analysis:

Undergraduate Level

Value	Label	Before Imputation		After Imputation	
		Frequency	Percent	Frequency	Percent
1	1st year/freshman	19,096	43.3	20,889	44.6
2	2nd year/sophomore	9,667	21.9	10,106	21.6
3	3rd year/junior	6,985	15.8	7,183	15.4
4	4th year/senior	7,673	17.4	7,911	16.9
5	5th year or higher	651	1.5	699	1.5
-9	Missing	2,716	--	--	--
Total		46,788	100.0	46,788	100.0

Variable Name: LOCALRES
Description: Local Residence
Cases Eligible for Imputation: All cases without valid values

Number with valid responses:	57,304
Number eligible for imputation:	3,816
Number not eligible for imputation:	N/A
Total:	61,120

Pre-Analysis:

Cross-tabulations were run on the following variables to determine the best predictors:

GENDER MARITAL AGECAT INJURIS UGRDLVL1 PROGTYPC

Hotdeck Imputation:

Sort variables:

1. OFCON1
*2. MARITAL[a/]
3. INURIS
4. PROGTYPC
5. UGRDLVL1
6. SCHOOL
7. AGE

*No break
a/ If MARITAL is missing MARITAL is set to 1 (single).

Number imputed by Hotdeck imputation = 3,816

Post Analysis:

Local Residence

Value	Label	Before Imputation		After Imputation	
		Frequency	Percent	Frequency	Percent
1	Campus Housing	12,776	22.3	13,225	21.6
2	Off-Campus	33,115	57.8	35,414	57.9
3	With Parents	11,413	19.9	12,481	20.4
-9	Missing	3,816	--	--	--
Total		61,120	100.0	61,120	100.0

Variable Name: RACE
Description: Race/Ethnicity
Cases Eligible for Imputation: All cases without valid values

Number with valid responses:	56,763
Number eligible for imputation:	4,357
Number not eligible for imputation:	N/A
Total:	61,120

Pre-Analysis:

Cross-tabulations were run on the following variables to determine the best predictors:

OFCON1, SEX, AGECAT

Hotdeck Imputation:

Sort variables:

*1. OFCON1
2. SCHOOL
3. MAJRCODE
4. GENDER
5. AGE

*No break

Number imputed by Hotdeck imputation = 4,357

Post Analysis:

Race

Value	Label	Before Imputation		After Imputation	
		Frequency	Percent	Frequency	Percent
1	American Indian	343	0.6	373	0.6
2	Asian	2,881	5.1	3,128	5.1
3	Black	5,085	9.0	5,782	9.5
4	Hispanic	4,193	7.4	4,734	7.7
5	White	44,261	78.0	41,103	77.1
-9	Missing	4,357	--	--	--
Total		61,120	100.0	61,120	100.0

Variable Name: INDEPINC
Description: Independent Student's/Spouse's 1988 AGI
Cases Eligible for Imputation: All cases without valid values

Number with valid responses:	18,902
Number eligible for imputation:	15,816
Number not eligible for imputation:	N/A
Total:	34,718

Regression Imputation:

For independent students in the student survey, ADJGRINC (S.S. 8.07 adjusted gross income 1988) were predicted using ALLINCOM (S.S. 8.05 estimated total income 1988), TOTINC (S.S. 5.15b and 5.16b, total income from work for student and spouse in 1988), and dummy variables indicating the control of the schools that the students attended (private, public, and proprietary). Students with incomes above $100,000 were excluded from the estimation procedure. Model R-square = 0.79, that is, almost 80 percent of the variation in ADJGRINC was explained by the predictor variables.

Number imputed by regression imputation = 9,314
Number remaining missing after regression imputation = 6,502

Hotdeck Imputation:

Cases eligible for hotdeck imputation = 6,502

Sort variables:

1. OFCON1
*2. PROGTYPC
3. SCHOOL
4. ATTEND
5. HSDEG
6. RACE
7. GENDER
8. MARITAL
9. AGE

* No break.

Note: Imputed incomes were not used as donors.

Number imputed by Hotdeck imputation = 6,502

Distribution of Independent Student's/Spouse's 1988 AGI

INDEPINC Distribution	Before Imputation	After Imputation
Mean	18,125	20,064
99 percent	85,000	80,055
95 percent	54,442	57,233
90 percent	42,000	46,000
75 percent	26,000	29,091
50 percent	12,243	14,651
25 percent	4,802	5,714
10 percent	1,000	1,897
5 percent	0	0
1 percent	0	0
N	18,902	34,718

Independent Student's 1988 Family AGI

DEPINC	Before Imputation		After Imputation	
	Frequency	Percent	Frequency	Percent
Less than 0	5	0.0	6	0.0
0	1,325	7.0	1,917	5.5
1-1k	567	3.0	712	2.1
1k-2k	655	3.5	1,030	3.0
2k-3k	772	4.1	1,312	3.8
3k-4k	836	4.4	1,433	4.1
4k-5k	774	4.1	1,355	3.9
5k-6k	766	4.1	1,320	3.8
6k-7k	696	3.7	1,241	3.6
7k-8k	690	3.7	1,166	3.4
8k-9k	634	3.4	1,038	3.0
9k-10k	646	3.4	1,124	3.2
10k-11k	473	2.5	812	2.3
11k-12k	540	2.9	968	2.8
12k-13k	440	2.3	762	2.2
13k-14k	409	2.2	737	2.1
14k-15k	453	2.4	805	2.3
15k-16k	390	2.1	706	2.0
16k-17k	367	1.9	730	2.1
17k-18k	385	2.0	713	2.1
18k-19k	329	1.7	635	1.8
19k-20k	383	2.0	732	2.1
20k-21k	256	1.4	498	1.4
21k-22k	317	1.7	595	1.7
22k-23k	265	1.4	564	1.6
23k-24k	316	1.7	584	1.7
24k-25k	317	1.7	626	1.8
25k-26k	260	1.4	540	1.6
26k-27k	256	1.4	521	1.5
27k-28k	229	1.2	451	1.3
28k-29k	176	0.9	381	1.1
29k-30k	338	1.8	609	1.8
30k-31k	138	0.7	325	0.9
31k-32k	210	1.1	424	1.2
32k-33k	137	0.7	310	0.9
33k-34k	148	0.8	313	0.9
34k-35k	219	1.2	448	1.3
35k-36k	140	0.7	320	0.9
36k-37k	132	0.7	263	0.8
37k-38k	166	0.9	340	1.0
38k-39k	83	0.4	196	0.6
39k-40k	199	1.1	418	1.2
40k-41k	79	0.4	195	0.6
41k-42k	100	0.5	227	0.7
42k-43k	93	0.5	236	0.7
43k-44k	63	0.3	183	0.5
44k-45k	134	0.7	282	0.8
45k-46k	71	0.4	167	0.5
46k-47k	70	0.4	175	0.5
47k-48k	93	0.5	223	0.6
48k-49k	64	0.3	150	0.4
49k-50k	153	0.8	294	0.8
Above 50k	1,145	6.1	2,606	7.5
Missing	15,816	--	--	--
Total	34,718	100.0	34,718	100.0

Variable Name: DEPINC
Description: Dependent Student's 1988 Family AGI
Cases Eligible for Imputation: All cases without valid values

Number with valid responses:	15,077
Number eligible for imputation:	11,155
Number not eligible for imputation:	N/A
Total:	26,232

Regression Imputation:

For respondents of the parent survey, ADJINC88 (P.S. 3.13 1988 adjusted gross income) were predicted using TOTINC89 (P.S. 3.05 1989 total income) and dummy variables indicating the control of the school that students attended (public, private, or proprietary). Parents with incomes below $100 or above $150,000 were excluded from the estimation procedure. Model R-square = 0.79, that is, about 80 percent of the variability in AGI was explained by the predictor variables.

Number imputed by regression imputation = 3,159
Number remaining missing after regression imputation = 7,996

Direct Imputation from Student Survey:

REFINC88 (derived from S.S. 7.04)

Number from student survey = 4,242
Number remaining missing = 3,754

Hotdeck Imputation:

Variable recoded:

PELLAMTC = 0 if Pell amount = 0,
1 if Pell amount = 1-299,
2 if Pell amount = 300-599,
3 if Pell amount = 600-899,
4 if Pell amount = 900-1,199,
5 if Pell amount = 1,200-1,500,
6 if Pell amount = 1,501-1,799,
7 if Pell amount = 1,800-2,099,
8 if Pell amount = 2,100-2,299,
9 if Pell amount = 2,300.

Sort variables:

*1. OFCON1
2. PELLAMTC
3. GINCEST (estimated AGI categories from parent survey)
4. DADEDUC
5. MOMEDUC

*No break

Note: Imputed incomes were not used as donors.

Number imputed by Hotdeck imputation = 3,754

Distribution of Dependent Student's 1988 Family AGI

DEPINC Distribution	Before Imputation	After Imputation
Mean	40,835	46,085
99 percent	200,000	200,000
95 percent	95,000	120,000
90 percent	75,000	85,000
75 percent	52,962	60,000
50 percent	34,508	37,820
25 percent	17,117	20,000
10 percent	4,288	6,000
5 percent	1,057	1,924
1 percent	0	0
N	15,077	26,232

Dependent Student's 1988 Family AGI

DEPINC	Before Imputation		After Imputation	
	Frequency	Percent	Frequency	Percent
Less than 0	14	0.1	14	0.1
0	536	3.6	790	3.0
1-1k	188	1.2	241	0.9
1k-2k	224	1.5	300	1.1
2k-3k	258	1.7	368	1.4
3k-4k	231	1.5	309	1.2
4k-5k	187	1.2	267	1.0
5k-6k	155	1.0	372	1.4
6k-7k	126	0.8	193	0.7
7k-8k	182	1.2	276	1.1
8k-9k	155	1.0	229	0.9
9k-10k	166	1.1	283	1.1
10k-11k	159	1.1	218	0.8
11k-12k	176	1.2	257	1.0
12k-13k	177	1.2	259	1.0
13k-14k	196	1.3	296	1.1
14k-15k	182	1.2	262	1.0
15k-16k	202	1.3	275	1.0
16k-17k	234	1.6	338	1.3
17k-18k	231	1.5	485	1.8
18k-19k	177	1.2	240	0.9
19k-20k	254	1.7	377	1.4
20k-21k	189	1.3	263	1.0
21k-22k	202	1.3	271	1.0
22k-23k	189	1.3	256	1.0
23k-24k	220	1.5	320	1.2
24k-25k	239	1.6	621	2.4
25k-26k	197	1.3	277	1.1
26k-27k	210	1.4	295	1.1
27k-28k	253	1.7	333	1.3
28k-29k	190	1.3	321	1.2
29k-30k	341	2.3	615	2.3
30k-31k	173	1.1	233	0.9
31k-32k	234	1.6	312	1.2
32k-33k	205	1.4	338	1.3
33k-34k	204	1.4	286	1.1
34k-35k	293	1.9	476	1.8
35k-36k	193	1.3	252	1.0
36k-37k	196	1.3	781	3.0
37k-38k	241	1.6	385	1.5
38k-39k	177	1.2	219	0.8
39k-40k	363	2.4	610	2.3
40k-41k	168	1.1	270	1.0
41k-42k	241	1.6	357	1.4
42k-43k	159	1.1	225	0.9
43k-44k	169	1.1	219	0.8
44k-45k	261	1.7	590	2.2
45k-46k	157	1.0	320	1.2
46k-47k	150	1.0	198	0.8
47k-48k	191	1.3	262	1.0
48k-49k	166	1.1	264	1.0
49k-50k	339	2.2	670	2.6
Above 50k	4,159	27.6	8,745	33.3
Missing	11,155	--	--	--
Total	26,232	100.0	26,232	100.0

Variable Name: SAI
Description: Student Aid Index
Cases Eligible for Imputation: All Pell recipients without valid values

Number with valid responses:	12,953
Number eligible for imputation:	300
Number not eligible for imputation:	N/A
Total:	13,253

Regression Imputation:

For Pell recipients, SAI were predicted using FAMINC (family income), PELLAMT (pell amount) and dummy variables indicating the control of school (public, private, and proprietary). Students with SAT great than $2,300 (n=11) were set to missing and new value imputed. Model R-square = 0.58, that is, almost 60 percent of the variation in SAI was explained by the predictor variables. Cases with negative predicted SAI were set to zero.

Number imputed by regression imputation = 300
Number remaining missing = 0

Frequency of Student Aid Index

SAI	Before Imputation		After Imputation	
	Frequency	Percent	Frequency	Percent
0	6,522	50.4	6,629	50.0
1-200	1,117	8.6	1,149	8.7
201-400	804	6.2	842	6.4
401-600	747	5.8	795	6.0
601-800	635	4.9	667	5.0
801-1,000	644	4.9	663	5.0
1,001-1,200	558	4.3	572	4.3
1,201-1,400	478	3.7	487	3.7
1,401-1,600	446	3.4	447	3.4
1,601-1,800	461	3.6	461	3.5
1,801-2,000	376	2.9	376	2.8
2,001-2,200	165	1.3	165	1.2
2,200+	11	--	--	--
Missing	289	--	--	--
Total	13,253	100.0	13,253	100.0

Variable Name: EFC
Description: Expected Family Contribution
Cases Eligible for Imputation: All cases without valid values

Number with valid responses:	28,720
Number eligible for imputation:	32,780
Number not eligible for imputation:	170
Total:	61,120

Regression Imputation:

Variables used to predict EFC were: FAMINC (family income), FAMVAL (farm value, missing were set to zero), FEDTAXES (federal taxes, if students were aided and FEDTAXES unknown, FEDTAXES were set to 0), FAMNUM (number in family, if missing FAMNUM were set to 3 for dependent students, or to 1 for independent students, POSTED (number in postsecondary, if missing set to 1), ANYAID (aided Yes/No), and depend (dependent or independent). The model $R^2 = 0.3$, that is, only about 30 percent of the variability in EFC was explained by the predictor variables. The plot of residuals against the observed values of EFC showed a linear trend suggesting a poor fit of the regression equation. We have also tried applying separate equations for aided versus unaided students and dependent versus independent students. None of these attempts appear to have a strong effect. Hot deck procedure was also tried using the same predictor variables as sort variables. This again was unsuccessful given the large number of missing.

To improve the EFC prediction, we recommend a more careful analyses of the predictor variables for influential observations, include other variables such as total asset, age of independent students, etc, to improve the predictive power of the equation, and to examine alternative procedure such as applying the congressional methodology formula for computing EFC. (NOTE: The EFC3 variable on the derived variable file represents a composite measure. If EFC was available from the student's records (EFC1), then this variable was used. If it was not available, ECF3 was based on the CM formula. For additional information, see the electronic codebook.

Appendix D

Standard Errors for Estimates in the Executive Summary

Summary Table 1

Standard Errors

Number of Students by Type and Control and Academic Level, in Thousands: 1989-90

Academic Level	All Institutions	Institutional Type and Control						
		Undergraduate					Graduate	
		Public Four-year	Public Two-year	Private Four-year	Private Two-year	Proprietary	Public	Private
Undergraduate	168	48	143	35	19	42	--	--
Graduate	25	--	--	--	--	--	23	13
All	174	48	143	35	19	42	23	13

Summary Table 2

Standard Errors

Number of Students, by Family Income and Academic Level, in Thousands: 1989-90

Academic Level	All Incomes	Family Income						
		Less than $10,000	$10,000-$20,000	$20,000-$30,000	$30,000-$40,000	$40,000-$50,000	$50,000-$100,000	$100,000 and over
Undergraduate	168	60	81	52	48	46	60	17
Graduate	25	9	7	11	8	7	8	1
All	174	58	79	56	48	47	60	16

Summary Table 3

Standard Errors

Number of Students, by Type and Control and Dependency Status, in Thousands: 1989-90

Students	All Institutions*	Institutional Type and Control						
		Undergraduate					Graduate	
		Public Four-year	Public Two-year	Private Four-year	Private Two-year	Proprietary	Public	Private
Dependent	114	46	88	37	9	28	3	2
Independent	158	53	119	39	14	30	22	12
All*	174	48	143	35	19	42	23	13

Summary Table 4

Standard Errors

Number of Students, by Family Income and Dependency Status, in Thousands: 1989-90

Students	All Incomes*	Family Income						
		Less than $10,000	$10,000-$20,000	$20,000-$30,000	$30,000-$40,000	$40,000-$50,000	$50,000-$100,000	$100,000 and over
Dependent	114	34	29	36	36	37	49	17
Independent	158	58	67	48	34	35	28	3
All*	174	58	79	56	48	47	60	16

* Since 65,500 weighted cases were unclassified, numbers do not add to totals.

Summary Table 5

Standard Errors
Percentage of Students Receiving Title IV Aid and Any Aid, by Type and Control and Academic Level: 1989-90

Academic Level		All Institutions	Institutional Type and Control						
			Undergraduate					Graduate	
			Public Four-year	Public Two-year	Private Four-year	Private Two-year	Proprietary	Public	Private
Undergraduate	percent Title IV aid	0.3	0.3	0.2	0.3	0.1	0.6	--	--
	percent any aid	0.4	0.5	0.7	0.3	0.2	0.4	--	--
Graduate	percent Title IV aid	0.6	--	--	--	--	--	1.7	1.7
	percent any aid	0.7	--	--	--	--	--	1.0	1.0
All	percent Title IV aid	0.3	0.3	0.2	0.3	0.1	0.5	0.1	0.2
	percent any aid	0.4	0.4	0.6	0.3	0.2	0.4	0.2	0.1

Summary Table 6

Standard Errors
Percentage of Students Receiving Title IV Aid and Any Aid, by Family Income and Academic Level: 1989-90

Academic Level		All Incomes	Family Income						
			Less than $10,000	$10,000-$20,000	$20,000-$30,000	$30,000-$40,000	$40,000-$50,000	$50,000-$100,000	$100,000 and over
Undergraduate	percent Title IV aid	0.3	0.8	0.4	0.3	0.3	0.2	0.2	0.0
	percent any aid	0.4	0.7	0.4	0.4	0.3	0.3	0.3	0.1
Graduate	percent Title IV aid	0.6	0.9	0.6	0.4	0.3	0.2	0.2	0.0
	percent any aid	0.7	0.5	0.4	0.3	0.3	0.3	0.3	0.1
All	percent Title IV aid	0.3	0.7	0.4	0.3	0.3	0.2	0.2	0.0
	percent any aid	0.4	0.6	0.3	0.3	0.2	0.2	0.3	0.0

Summary Table 7

Standard Errors

Percentage of Students Receiving Title IV Aid and Any Aid, by Type and Control and Dependency Status: 1989-90

Students		All Institutions	Institutional Type and Control						
			Undergraduate					Graduate	
			Public Four-year	Public Two-year	Private Four-year	Private Two-year	Proprietary	Public	Private
Dependent	percent Title IV aid	0.5	0.8	0.8	0.8	0.2	0.8	0.0	0.1
	percent any aid	0.7	0.8	0.8	0.7	0.2	0.6	0.1	0.1
Independent	percent Title IV aid	0.4	0.5	0.7	0.4	0.1	0.6	0.2	0.4
	percent any aid	0.4	0.4	0.9	0.5	0.2	0.5	0.3	0.3
All	percent Title IV aid	0.3	0.3	0.2	0.3	0.1	0.5	0.1	0.2
	percent any aid	0.4	0.4	0.6	0.3	0.2	0.4	0.2	0.1

Summary Table 8

Standard Errors

Percentage of Students Receiving Title IV Aid and Any Aid, by Family Income and Dependency Status: 1989-90

Students		All Incomes	Family Income						
			Less than $10,000	$10,000-$20,000	$20,000-$30,000	$30,000-$40,000	$40,000-$50,000	$50,000-$100,000	$100,000 and over
Dependent	percent Title IV aid	0.5	1.0	0.5	0.5	0.5	0.3	0.4	0.1
	percent any aid	0.7	0.7	0.4	0.5	0.4	0.3	0.5	0.1
Independent	percent Title IV aid	0.4	0.9	0.6	0.5	0.2	0.1	0.1	0.0
	percent any aid	0.4	0.8	0.5	0.5	0.3	0.2	0.2	0.0
All	percent Title IV aid	0.3	0.7	0.4	0.3	0.3	0.2	0.2	0.0
	percent any aid	0.4	0.6	0.3	0.3	0.2	0.2	0.3	0.0

Summary Table 9

Standard Errors

Average Title IV and Total Aid of Students, by Type and Control and Academic Level: 1989-90

Academic Level		All Institutions	Institutional Type and Control						
			Undergraduate					Graduate	
			Public Four-year	Public Two-year	Private Four-year	Private Two-year	Proprietary	Public	Private
Undergraduate	average Title IV aid	$ 40	$37	$54	$ 76	$177	$109	--	--
	average total aid	47	45	71	161	168	98	--	--
Graduate	average Title IV aid	87	--	--	--	--	--	$101	$105
	average total aid	221	--	--	--	--	--	185	358
All	average Title IV aid	43	37	54	76	177	109	101	105
	average total aid	60	45	71	161	168	98	185	358

Summary Table 10

Standard Errors

Average Title IV and Total Aid of Students, by Family Income and Academic Level: 1989-90

Students		All Incomes	Family Income						
			Less than $10,000	$10,000-$20,000	$20,000-$30,000	$30,000-$40,000	$40,000-$50,000	$50,000-$100,000	$100,000 and over
Undergraduate	average Title IV aid	$ 40	$ 59	$ 55	$ 56	$ 55	$ 77	$ 63	$250
	average total aid	47	67	54	92	101	115	72	181
Graduate	average Title IV aid	87	122	123	106	170	185	169	367
	average total aid	221	265	261	228	282	187	169	708
All	average Title IV aid	43	63	58	52	54	74	65	237
	average total aid	60	83	83	88	95	105	70	175

Summary Table 11

Standard Errors

Average Title IV and Total Aid of Students, by Type and Control and Dependency Status: 1989-90

Students		All Institutions	Institutional Type and Control						
			Undergraduate					Graduate	
			Public Four-year	Public Two-year	Private Four-year	Private Two-year	Proprietary	Public	Private
Dependent	average Title IV aid	$41	$35	$ 61	$ 66	$182	$165	$343	$251
	average total aid	80	45	157	176	145	132	281	423
Independent	average Title IV aid	57	54	72	123	315	121	96	100
	average total aid	80	67	92	181	264	117	192	362
All	average Title IV aid	43	37	54	76	177	109	101	105
	average total aid	60	45	71	161	168	98	185	358

Summary Table 12

Standard Errors

Average Title IV and Total Aid of Students, by Family Income and Dependency Status: 1989-90

Students		All Incomes	Family Income						
			Less than $10,000	$10,000-$20,000	$20,000-$30,000	$30,000-$40,000	$40,000-$50,000	$50,000-$100,000	$100,000 and over
Dependent	average Title IV aid	$41	$71	$ 96	$ 58	$ 52	$ 85	$ 60	$280
	average total aid	80	43	103	118	139	120	82	166
Independent	average Title IV aid	57	72	67	92	122	184	193	low N
	average total aid	80	88	110	113	135	259	124	816
All	average Title IV aid	43	63	58	52	54	74	65	237
	average total aid	60	83	83	88	95	105	70	175

Appendix E

Memos on Weighting Procedures and Description of 1987 NPSAS sample

WESTAT
An Employee-Owned Research Corporation
1650 Research Blvd. • Rockville, MD 20850-3129 • 301 251-1500 • Fax 301 294-2034

MEMORANDUM

TO: Drew Malizio November 26, 1990

FROM: Mike Brick

SUBJECT: Preliminary Abstract File Weight Specifications

This memorandum contains the weighting and variance procedures proposed for the preliminary student abstract data file to be delivered in December 1990. We would also like to determine if the format of these specifications are at the level you desire for future weight and variance specifications for NPSAS. The final section raises some issues about the number and types of weights that we will be producing for the final deliverable data files.

I. Weighting for Institutions

The institutions were sampled for NPSAS in a two stage sample. The first stage was the sampling of primary sampling units (PSU's) composed of geographic areas. The probability of selection for a PSU is denoted by P_{hi}, where h identifies the primary sampling unit stratum and i identifies the PSU within a stratum. These probabilities were identical to those defined in the 1987 NPSAS and do not require any recomputation. In 1990, a new PSU which included all of Puerto Rico was added. Since it was added as a self-representing PSU, its probability is unity. In all, there are 121 PSU's in NPSAS '90, 51 are self-representing (SR) and 70 are non-self-representing (NSR).

The second stage of sampling institutions was the selection of units from within the sampled PSU's. Different procedures were followed for institutions with very large enrollments and for some schools sampled from an supplemental list. These exceptions will be described below.

The sampling procedure within PSU's was to stratify the institutions, allocate a sample size to each stratum, assign a measure of size to each institution, and then

select a systematic sample with probabilities proportionate to their measure of size from each stratum across the PSU's. Suppressing the previous subscripts for ease of presentation, the probability of selection at this stage is equal to

$$P_{jk} = \frac{n_j \, MOS_{jk}}{\sum MOS_{jk}} ,$$

where j refers to the institutional stratum and k refers to the specific institution within stratum j. This sampling resulted in the selection of 114 self-representing institutions (from SR PSU's) and 1,194 non-self-representing institutions (either NSR institutions from SR PSU's or SR or NSR institutions from NSR PSU's).

The institutions with the very largest enrollments were sampled with certainty prior to the selection of the primary sampling units. These units have probabilities of sampling equal to unity, so that P_{hi} and P_{jk} are both assigned to be equal to unity for purposes of weighting. There were 225 of these institutions sampled. This brought the total number of sampled institutions to 1,533.

A supplemental sample of institutions on the Stafford Loan Institution file was selected in order to ensure the coverage of institutions with small enrollments of short duration. The sample was selected entirely within the 121 sampled NPSAS PSU's. An equal probability sample of 476 of these institutions was selected (since enrollment was not on the list, an enrollment related measure of size for each institution could not be obtained). The sampled institutions were then compared to the frame used for basic NPSAS sample and duplicates removed. Of the 476, only 81 were released to the field for clarification and further unduplication. The probabilities of selection for these institutions can, therefore, be represented in the same fashion and notation as the basic sample.

The basic (unadjusted, inflation) weight for an institution is then the inverse of its probability of selection, or

$$I_{hijk} = (P_{hi} \times P_{jk})^{-1} ,$$

where the subscripts are as defined above. This weight does not account for the fact that some institutions failed to participate in the survey. To accommodate this, a nonresponse adjustment at the institution level is needed.

To adjust for institution level nonresponse a status flag will be created for each institution indicating if it was eligible and participating, eligible and not participating, or ineligible. The nonresponse adjustment factor will be the sum of the enrollment (defined at the point of sampling) in the eligible institutions divided by the sum of the enrollment in the eligible and participating institutions. The nonresponse adjustment factors will be constructed separately in classes defined by level and control of institution. The institution level and control classes are:

1 -public-less than 2 year,
2 -public-2 year,
3 -public-other 4 year,
4 -public-PhD or first professional,
5 -private not for profit-less than 2 year,
6 -private not for profit-2 year,
7 -private not for profit-other 4 year,
8 -private not for profit-PhD or first professional,
9 -private for profit-less than 2 year,
10 -private for profit-2 year or greater.

The adjustment factors can be written as

$$NRF_m = \frac{\sum_{k \text{ is eligible}} E_{mk}}{\sum_{\substack{k \text{ is eligible} \\ \text{and participating}}} E_{mk}} ,$$

where E is the enrollment (if unknown a value will be imputed) and m is the level and control of the institution, as defined above. The nonresponse adjusted weight for an institution is the product of the basic weight and the appropriate nonresponse adjustment factor,

$$I'_{hijkm} = I_{hijkm} \times NRF_m .$$

This weight is valid for the sampling of institutions for the October data collection, since all institutions were included in that effort. However, the nonfall data collection efforts (August, February, and June) were restricted to subsamples of the institutions selected for October. Also the institutions which participated in the fall data

collection may not have participated in all of the nonfall efforts. These and similar considerations must be addressed for the weighting of the nonfall samples of institutions.

The previous formulations for the weights associated with sampling institutions can be extended slightly to account for the nonfall subsampling. If we denoted the time period by the subscript t, then the basic weight can be rewritten as

$$I_{thijk} = (P_{hi} \times P_{tjk})^{-1} \ ,$$

where the P_{tjk} incorporates the additional subsampling done for nonfall time period t. Since only less than 2 year institutions were sampled for the June data collection, there will be a bias in the estimates arising from the failure to cover 2 and 4 year institutions during this period.

The nonresponse adjustment must also specifically account for the time period. Therefore, the status flags will indicate for each institution if it was eligible and participating, eligible and not participating, or ineligible by time period. It should be noted eligibility could change during the year. For example, some institutions were open for part of the year but later closed becoming ineligible. Another consideration is that some institutions had only one registration period in the year, so that no lists of nonfall students were present. These institutions are considered eligible and participating. This designation is necessary in order to develop proper nonresponse adjustment factors.

The nonresponse adjustment factors can be written as before, with the addition of a subscript for data collection periods,

$$NRF_{tm} = \frac{\sum_{\substack{k \text{ is eligible} \\ \text{for period t}}} E_{tmk}}{\sum_{\substack{k \text{ is eligible} \\ \text{and participating} \\ \text{in period t}}} E_{tmk}} \ .$$

Note that E is still the enrollment at the time of sampling (not a nonfall estimate of enrollment) because that is not available for most institutions which did not participate. The nonresponse adjusted institution weight which accounts for the sampling at different time periods can be written as

$$I''_{thijkm} = I_{thijkm} \times NRF_{tm} .$$

This is the weight that will be associated with an institution for a particular sampling period.

II. Weighting for Students

The weights for making estimates at the student level are the product of the institution weight and the weight within an institution. The weights that account for sampling within an institution are described first, and then the specifications for the overall weights are given.

The sampling of students within a sampled institution was a stratified systematic sample in which all students in an institution and stratum had equal probabilities of selection. There were up to three student strata in an institution corresponding to the level of the student (first professional, graduate, and all others). The sampling was also done separately for each of the four data collection time periods. The within institution level weight for a student can be written as

$$S_{tkps} = \frac{L_{tkp}}{n_{tkp}} ,$$

where L_{tkp} is the number of students on the enrollment list from institution k, at time period t, and student level p, n_{tkp} is the number of students sampled from that list, and s identifies the student sampled from the list. Note that the subscripts associated with an institution, other than k, have been suppressed to avoid confusion.

This formulation would be sufficient if the students on the lists from the same institutions for different time periods were mutually exclusive. Of course, this was not the case in practice. In most of the institutions, some students appeared on the lists from more than one time period. To permit students only one chance of being in the sample, the students selected in the nonfall periods were checked against the enrollment lists obtained from the other time periods and eliminated from the sample if they had already had some chance of being selected. The weight can then be written as

$$S'_{tkps} = \frac{L_{tkp}}{n_{tkp}} \delta(s) ,$$

where $\delta(s)$ is zero if student s was previously listed, and unity otherwise. Note that for the fall data collection the final term is unnecessary, since the fall enrollment list was the base period.

The unduplication of students sampled for nonfall periods against the lists was a difficult procedure that resulted in some errors. In particular, some students who were sampled in a nonfall period and appeared on a previous enrollment list were not identified and dropped from the sample. This happened most frequently when the enrollment lists for the different time periods were in different order, the student moved on to graduate school, or the student changed names. Even though the percent of errors in the unduplication process was relatively small (probably less than 1 percent), these errors could introduce small biases in the estimates unless adjustments are made.

The proposed adjustment consists of correcting the probability of selection for the student from the institution and including the student in the sample. An alternative approach is to eliminate the student from the sample, but this approach results in the loss of all the data collected on that student. This loss is not necessary since the adjustment of the probability of selection is simple and should nearly eliminate any biases.

The weight for a student who was selected as a nonfall student, even though listed on a previous enrollment list from the institution, will be divided by a factor of two to account for these students have two chances of being sampled (if students could have been sampled three times, then the factor should be three instead of two). The within institution weight with this adjustment can be written as

$$S'_{tkps} = \frac{L_{tkp}}{n_{tkp}} \frac{\delta(s)}{2} \quad \text{if student was eligible for sampling twice,}$$

$$S'_{tkps} = \frac{L_{tkp}}{n_{tkp}} \delta(s) \quad \text{otherwise}$$

The student weight is then the product of the nonresponse adjusted institution weight and the appropriate student weight. This weight is

$$W_{thijkmps} = I''_{thijkm} \times S'_{thijkmps}$$

This is the base student weight with institution level nonresponse adjustments. The particular weights that will be produced for making estimates from NPSAS will be based on this weight.

III. Preliminary Abstract Student Weights

The weights for the preliminary file delivered in December 1990 will consist of the base student weight with institution level nonresponse adjustments, as given above. Even though the preliminary data file will be restricted to aided, undergraduate students all the preceding steps will still apply. No further student level nonresponse adjustment will be included. Furthermore, if the examination of the errors in the unduplication process is not yet be completed, the adjustment described above for multiple probabilities of selection will not be incorporated in the preliminary weight.

The preliminary file will include any aided, undergraduate students with a record abstract. The abstract could have been collected as a part of the fall data collection effort, or later during the updating period. Therefore, the definition for a participating institution for a time period must depend on the collection of information on the student. An institution will be considered participating for the time period if there are any record abstracts for students sampled during that period. If the institution was included in the sample for the time period but no students were sampled (all duplicates, etc.,), then the institution will also be considered to be participating.

One concern in producing estimates from NPSAS '90 has been the potential for biases since students may enroll in more than one institution during the year, and thus be eligible for sampling more than once. This has been called the 'transfer' student problem.

For the preliminary file, and all other files that only include data collected from the institution, the transfers are not a problem. The reason is that the institution provides complete year data on the activities of the student *at that institution*, but it does not have access to data about the student at any other institution. In order to produce unbiased national estimates of characteristics which are complete and consistent with program totals, the weighting procedure described above is appropriate. Characteristics, such as the

number of students with Pell grants and the dollar amount of those grants, would be underestimated if transfer students were eliminated. We have called this the student per institution approach and discussed its properties in previous memoranda.

IV. Variance Estimation Procedures for the Preliminary File

It is our understanding that a specialized variance estimation program created by NCES will be used to produce estimates of the sampling errors for the estimates from the preliminary file. This program uses Taylor Series expansion methods to approximate the sampling errors for the statistics. Therefore, we will produce the control variables needed for this program. Since we do not have experience with this particular program, the specifications below are based on some assumptions and discussions. If these are incorrect, we will modify them to satisfy the program.

Two key variables, STRATUM and PSU will be created and encoded on each student abstract record that has a weight. The variable STRATUM will be equal to the OFCON variable in the preliminary file; it is the same as the institution level and control variable used in the nonresponse adjustments described in section I.

The PSU variable will be an integer constructed based primarily on the institution from which the student was sampled. Each NSR institution will be designated as a PSU. For SR institutions, two PSU's will be created and the sampled students will be randomly divided into those PSU's. We assume that the number of PSU's per stratum is not restricted so that either even or odd numbers are acceptable.

An additional procedure may be required since the variance estimation program requires at least 4 records per STRATUM and PSU. In virtually all institutions, at least 4 students were sampled and record abstracts were completed. However, the preliminary file will only contain data on the aided, undergraduate students. This subsetting will probably reduce the number of students in the file to less than one-third of all the students sampled, and result in more PSU's having 4 or fewer units.

To prevent this subsetting from having adverse effects on the ability to run the variance estimation program, we propose to collapse institutions so that at least 4 students are in each STRATUM-PSU. The collapsing will be done within STRATUM by sorting all institutions by their basic institution level weight and assigning any institution

with fewer than 4 records to an adjacent PSU. The PSU numbers will then be reassigned so that they are consecutive within a STRATUM.

The variance estimation procedures for the final files will be quite different from these since they will use replication methods. The replication variances will incorporate the full sample design used in NPSAS '90 and will not be limited to the data from the preliminary file.

V. Final Weights

The final files will be accompanied by estimation weights and replicate weights that will go beyond the procedures outlined for the preliminary file. For example, adjustments for student level nonresponse, and ratio and poststratification adjustments will be considered for the final weights. The replicate variance procedures are not even discussed above. Disclosure avoidance techniques will also be incorporated to insure that institutions can not be identified from the student level weights.

In addition to these technical details, a basic issue is the number and types of weights that are needed. Weights for record abstracts, student questionnaires and parent questionnaires have been discussed and need to be clarified. Below we discuss our current understanding of the weights for final files and some alternatives that could be considered.

The record abstract weights for the final files are only a slight extension of the weights described above for the preliminary file. The primary additions will be nonresponse adjustments and ratio and/or poststratification adjustments. Nonresponse adjustments will be necessary because not all abstracts were obtained in the two data collection efforts, including some for which nonfall abstracts were not included by design. We currently plan to consider these to be incomplete data (regardless of whether a student questionnaire was completed) and make nonresponse adjustments for the abstract file.

Ratio adjustments can be made to lower the variance associated with sampling institutions. These adjustments can be made from existing data in IPEDS. If poststratification to the number of students awarded Pell grants or the dollar amount of Pell grants is desired, then the appropriate files from the Pell grants program must be made available. The consistency of the data on the file used for poststratification and the NPSAS file will also have to be verified in order to prevent biases in the estimates.

The student questionnaire data will also be weighted so that national estimates can be derived for these data. The student questionnaire totals can be adjusted to equal the number of students from the abstract file by adjustment cell class (for example, by level and control of institution and level of student). Prior to making this adjustment, student level nonresponse adjustments are necessary. NCES mentioned considering the idea of labelling those students for whom there is a great deal of record abstract data but did not respond to the student questionnaire as completes for this file. If this is done, then the algorithm for designating them as completed needs to be established.

The other significant issue associated with the student questionnaire weighting is the problem of handling 'transfer' students. In general, the issue is whether to produce national estimates of students (these would not be consistent with any of the federal program data or IPEDS) or students per institution. The simplest solution is to form two weights, one for each type of estimate. Of course, this solution may not be practical for users, since they must decided on which weights to use for different purposes.

Since so few students attend more than one institution in a year, a reasonable compromise might be to produce student per institution weights for public use and estimate the number and characteristics of students who attend more than one institution per year. These students can also be flagged on the data file. This solution still requires that both sets of weights be developed.

Parent weights, really weights for students whose parents were sampled for interviews, are directly tied to the student weights but differ significantly from the other weights. The parent interview was intended to complement the data collected in the student and record abstract efforts. The universe of parents (we'll use this term loosely) is defined to include only parts of the undergraduate (less than 30 year old), student population. In fact, the only two inference population which appear to be of any separate interest are the dependent, unaided population and the first-time, first-year population. Both of these populations were included with certainty in the parent survey, but only if the student completed the student questionnaire.

Because of its limited usefulness for producing estimates of interest (the primary objective of the parent survey - to provide more accurate financial data for the student to be used when abstract data were not available - is not diminished by the fact that

it is restricted in this fashion), we propose simple estimation procedures. The student weight for those responding to the parent interview will be adjusted to account for nonresponse. This nonresponse adjustment will be done in cells, at least the cells corresponding to the groups of parents sampled. If special weights for any particular group of parents is desired, we will consider other estimation methods to meet these requirements.

WESTAT
An Employee-Owned Research Corporation

MEMORANDUM

TO: Drew Malizio May 7, 1991

FROM: Mike Brick

SUBJECT: NPSAS Weights File

Attached are two flat files with weights and assorted variables for producing estimates from the NPSAS:90 files. Documentation of the variables in each file is also attached. The contents of the files are summarized below.

The first file is ABSWGT.DAT and it contains variables appropriate for use with the file of respondents to the record abstract data collection (68,599 respondents). The ABSWGT variable is the weight developed using procedures similar to those described in my November 26, 1990 memo to you. It is a student per institution weight, as all abstract weights must be. The institution weights, a first step in the process of creating these weights, were ratioed to IPEDS totals (rather than using a simple nonresponse adjustment implied in the memo). The other notable difference was the inclusion of a undercoverage adjustment in 3 of the strata (both for profit strata and the less than 2 year nonprofit stratum) for the nonfall students. This adjustment was premised on evidence of some lack of coverage of students within these institutions for periods outside of the fall. The adjustment factor was developed based on counts of Pell recipients in the strata.

The file contains a second weight, PSABSWT, that was developed in accordance to NCES specifications. The basic weight described above was the starting point for this weight. The weights for the students enrolled in the fall were poststratified to counts from IPEDS. The undercoverage adjustment for the nonfall students described above was also implemented. A final adjustment was made to the weights of those students with Pell awards. The weights for Pell recipients were adjusted so that the total number of recipients by level and control of institution and Pell award amount were equal to the number of recipients reported in a draft of the Pell End of Year Report for 1989-90.

Two other variables are provided for use with Taylor series expansion programs for computing variances. These are the OFCON (the stratum variable in the program) and the PSU. Note that although stratum is the same regardless of which file is used, the PSU is specific to the abstract file since each PSU was constrained to have at least 4 respondents.

The second file is called STUWGT.DAT and it contains the variables for use with the file of respondents that were identified by NCES using record abstract and student CATI data files (61,120). There are three weights on this file, two are student per institution weights and one is at the student level. The KEEPWT is a student per institution weight that was developed along the same lines as the ABSWGT, e.g., it reflects the inverse of the probability of selection, nonresponse adjustments for sampled students which were not included in the file, and an adjustment for undercoverage in 3 strata. The other student per institution weight is called PSKEEPWT. It was developed along the lines of the PSABSWT, according to the same NCES specifications. It involved poststratification of the fall students to IPEDS totals and adjustments to Pell recipients counts by type and award amount.

The third weight is called PSSTUWT and it is the only student level weight in the files. The weight was created by dividing the PSKEEPWT by two if the student was attended two or more institutions during the 1989-90 year, unless the student was sampled twice at different institutions. The purpose of the adjustment is to provide estimates of unduplicated counts of enrolled students, even if they attended more than one institution. This type of weight would be appropriate for unduplicated counts of students and for estimating characteristics that are totaled over all institutions a student attends (the abstract data do not fit this requirement).

Filename: ABSWGT.DAT
RECFM=F LRECL=48

Records: 68,599

Column	Variable	Format	Description
1	STU_ID	$char13.	Student ID
15	NCESID	6.	NCES Student ID
22	OFCON	2.	STRATUM (Control & Level)
25	FALL_FLG	$char1.	Sampling Flag 1=Fall 2=Non-Fall
27	ABSWGT	8.2	Abstract Weight
36	PSABSWT	8.2	Post Stratified Abstract Weight
45	PSU	4.	Abstract PSU

Filename: STUWGT.DAT
RECFM=F LRECL=57

Records: 61,120

Column	Variable	Format	Description
1	STU_ID	$char13.	Student ID
15	NCESID	6.	NCES Student ID
22	OFCON	2.	STRATUM (Control & Level)
25	FALL_FLG	$char1.	Sampling Flag 1=Fall 2=Non-Fall
27	KEEPWT	8.2	Keep Weight
36	PSKEEPWT	8.2	Post Stratified Keep Weight
45	PSSTUWT	8.2	Post Stratified Student Weight
54	PSU	4.	Keep PSU

[Reproduced from the NPSAS:87 Methodology Report, NCES 90-309, pp. 9-11]

3. 1987 NPSAS IN-SCHOOL SAMPLE

3.1 - Sample Design

The sample design for the 1987 NPSAS in-school sample involved three stages of sampling-area sampling, institution sampling, and student sampling. A subsample of the in-school sample of students was used for the parent survey.

Area sampling was used to reduce field data collection costs. Because of the lack of a complete national listing of all postsecondary students for a sample frame, it was necessary to sample students at institutions. Within three-stage design, students for the 1987 NPSAS in-school sample were selected from sammpled institutions that were located within sampled geographic closters (areas).
The overall purposes of the 1987 NPSAS in-school sample design were to adequately represent students in all sectors of postsecondary education, to minimize the variability of the estimates of characteristics of the students for selected domains, and to allow economical field data collection and data processing,

3.2 Area sampling

Area sampling has the potential of increasing the variability of national estimates. The impact of this was minimized by stratification, assigning differential probabilities of selection, and controlling the number and definition of clusters.

The rationale behind area sampling was to reduce field data collection travel costs to sampled institutions and to students and parents for nonrespondent followup interviews. The 1985-1986 Integrated Postsecondary Education Data System (IPEDS) postsecondary universe file, however, was not available for use as a sample frame as planned. Area sampling was necessary to ensure coverage of all institutional segments through the development of a listing of all postsecondary institutions within sampled areas.

Because of the use of area sampling, it was not necessary to have a complete national postsecondary universe file. However, a national file was needed for drawing the very largest schools across the country and for assigning measures of size to areas on the area sample frame. For this purpose, a preliminary universe file was created by combining and unduplicating the 1984-1985 Higher Education General Information Survey (HEGIS) XIX Opening Fall Enrollment File and the 1983-1984 Pell/Campus-Based Institution file. These files provided a national list of all accredited and Pell-eligible postsecondary institutions with information on their type, control, enrollment, and address. Institutions were stratified by type, and control and any institution which exceeded a cut off enrollment in its type was selected with certainty. The cut-off point for certainty institutions was set equal to one-half of the sampling interval for the stratum. These factors were based on the initial estimates of the number of institutions needed in each stratum (type and control) and the size (enrollment of that stratum). A total of 162 institutions were selected with certainty in this step and were removed from the preliminary institution file before area sampling.
The preliminary institution universe file without certainty schools contained 6,387 schools.

The next step in the process was to identify all three-digit zip code areas in all fifty States and the District of Columbia for use as the area sample clusters. A minimum size of seven postsecondary institutions and 1,000 students for each cluster was defined. If a three digit zip code areas did not meet this minimum size requirement then it was clustered with other adjacent three-digit sip code areas to form a Primary Sampling Unit (PSU). Clusters did not cross state boundaries. A universe of 361 PSUs was created in this manner.

Some of the Primary Sampling Units were very large in counts of institutions and/or enrollment. The next step of area sampling was to select the largest PSUs with certainty. A PSU was selected with certainty if the total enrollment in the PSU exceeded one-half the sampling interval. The preliminary universe file was used to determine these conditions. A total sample of 120 PSUs was desired. Of the 361 PSUs in the universe, 50 were large enough to be included with certainty.

A sample of 70 Primary Sampling Units from the remaining 311 noncertainty PSUs was necessary in order to provide the desired 120 sample PSUs. Each PSU was assigned a measure of size that depended upon the total number of students in the PSU and the number of students in four different types of institutions. A function of these numbers was used as the measure of size for each PSU. This measure of size was used instead of total enrollment to ensure adequate representation of smaller specialized institutions.

The next step was to stratify the PSUs. The primary stratification variable at this stage of sampling was the state in which the PSU was located. The strata were designed so that each stratum had roughly the same total size, where total size was equal to the sum of the measures of size of all the PSUs in the stratum. If the PSUs in a state were not large enough to constitute an entire stratum (or were so large that they were greater than one stratum but less than two), then PSUs from different states were placed in the same stratum based upon indices reflecting the nature of the level of state aid to postsecondary education. In all, 35 strata were formed, each roughly of equal size.

The PSUs were then assigned a probability of selection proportional to their measure of size. The PSUs were sorted in a stratum by state and within state by their measure of size.

The PSUs were sampled systematically from each stratum with probability proportional to their measure of size. The final area sample consisted of 120 PSUs. Fifty PSUs were selected with certainty and 70 PSUs were selected with probability proportional to their measure of size. With the exception of four states, each state contained at least one PSU. Each PSU consists of one or more contiguous three-digit zip code areas.

[Additional information (not in the 1987 Methodology Report) on the average size, standard deviation, minimum and maximums of the self-representing (certainty PSUs) and other PSUs is shown below.]

	AVERAGE SIZE	STANDARD DEVIATION	FREQ	MIN	MAX
All PSUs	48820.32	58368.50	121	4894	567548
Certainty PSUs	82389.38	78252.89	51	45103	567548
Others PSUs	24842.43	11278.35	70	4894	43380

United States
Department of Education
Washington, D.C. 20208–5652

Official Business
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G–17

FOURTH CLASS BOOK RATE